IN THE UNITED STATES PATENT OFFICE

In re patent application of:	)	Before the Examiner:
	)	Andrew R. Graham
Albert S. Feng et al.	)	
	)	
Application No. 09/193,058	)	
	)	Group Art Unit: 2644
Filed: November 16, 1998	)	
	)	
BINAURAL SIGNAL PROCESSING TECHNIQUES	)	
	)	

DECLARATION UNDER 37 CFR §1.131

We, the inventors of the above-indicated patent application, hereby declare as follows:

1. We have reviewed U.S. Patent Application No. 09/193,058 (the "'Subject Application"), which is attached in exhibit A as originally filed on November 16, 1998, and U.S. Patent Application No. 08/666,757, (the "Parent Application"), which is attached in exhibit B as filed June 19, 1996. The Parent Application issued on April 24, 2001 as U.S. Patent No. 6,222,927.

2. We have each reviewed claims 34, 35, 37-42, 44-48, 50-54, and 56-66 that have been or will be proposed for the Subject Application (the "Patent Claims"), a copy of which is attached as exhibit C.

3. We have compared the Patent Claims of exhibit C to the text of the Parent Application of exhibit B and the text of the Subject Application of Exhibit C. Based on this comparison, the Parent Application describes in writing the inventions defined by the Patent Claims in

the text common to the Subject Application. Specifically, the written description of the Parent Application appearing on pages 1-3 is substantially the same as the written description of the Subject Application appearing on page 1, line 13 – page 3, line 17; the written description of the Parent Application appearing on pages 8-17 is substantially the same as the written description of the Subject Application appearing on page 9 – page 17, line 27; and the written description of the Parent Application appearing on pages 18-19 (Experimental Section) is substantially the same as the written description of the Subject Application appearing on page 37 (Example One).

4. Research directed to extracting a desired acoustic signal from a noisy acoustic environment, such as that associated with the "cocktail party effect," began prior to September 1995 and continued at least into the year 1996 as evidenced by section 14 (pages 6-7) of the attached exhibit D. Exhibit D includes an Invention Disclosure Report prepared by the inventors, a cover letter dated May 9, 1996, and supplemental information for the Invention Disclosure Report, and has been partially redacted.

5. Based on information and belief, the cover letter dated May 9, 1996 forwarded the Invention Disclosure Report to the law firm that prepared and filed the Parent Application and the Subject Application. Such firm subsequently received the supplemental information of exhibit D before June 1996.

6. Based on information and belief, at least some of the records and/or files referenced in exhibit D (including that referenced in section 14(c)) cannot be located due to inadvertent

destruction or loss in connection with one or more equipment upgrades in years subsequent to the filing of the Parent Application.

7. The subtractive processing algorithm, accompanying mathematical formulae, and other aspects set forth in exhibit D were established at least as early as September 15, 1995. No later than this date, we formed in our minds a definite and permanent idea of the complete and operative inventions defined by the methods of the Patent Claims with the establishment of these concepts. We confirm our recollection of this timing by its chronological relationship to the information set forth in the documents of Exhibit E that all existed prior to September 1995.

8. From before September 18, 1995 through filing of the Parent Application on June 19, 1996, the inventors have diligently continued research, development, evaluation, and experimentation regarding the inventions defined by the Patent Claims. Such activity before September 18, 1995 is supported by at least the information set forth in exhibit D. Also, in exhibit D note the formation of one corresponding research team August of 1994 (section 14(a) followed by an initial written record September of 1995 (section 14 (b)). This initial record was prepared by inventor Chen Lui shortly after joining the research effort during late August of 1995 as corroborated by certain entries of exhibit E attached.

9. After September 18, 1995, activities continued with the preparation of a detailed research initiative proposal dated November 29, 1995, in which the subject matter of exhibit E corresponds to the text with the heading "aim #1." A copy of a draft of this proposal is

provided in exhibit F.

10. Before February 9, 1996, experimental activities included computer simulation of the processes described in at least independent claims 34, 46, 61, and 62 of the Patent Claims, as explained in section 7 (pages 4-5) of the Invention Disclosure Record of exhibit D. These efforts are also discussed in the communication to inventor, Dr. Albert Feng, on February 12, 1996 from Lynn Huerta, which is included as exhibit G. To those skilled in the art to which the inventions of the Patent Claims pertain, this type of simulation establishes performance of the corresponding inventions in the intended manner.

11. Before forwarding the Invention Disclosure Report of exhibit D to counsel on or about May 9, 1996, testing of an experimental prototype was conducted further establishing performance of the inventions of at least independent claims 34, 46, 61, and 62 of the Patent Claims in the intended manner. This experimentation is further detailed in section 10 of the Invention Disclosure Report of exhibit D. The experimental section of the Parent Application (pages 18-19) and the Subject Application (page 37) correspond to our experimental activities.

12. Based on information and belief, counsel that received the Invention Disclosure Report of exhibit D promptly reviewed it and arranged an interview by telephone with at least one of the inventors, Dr. Albert Feng, to further discuss the information contained therein. The supplement to the Invention Disclosure Report of exhibit D was subsequently

received by such counsel before June 1996.

13. Based on information and belief, counsel reviewed the Invention Disclosure Report and prepared the Parent Application from the materials provided between May 9, 1996 and its filing date of June 19, 1996.

14. The undersigned, being hereby warned that willful false statements and the like are punishable by fine or imprisonment, or both (18 U. S C. 1001), and may jeopardize the validity of the application or any patent issuing thereon, declares that all statements made of his/her own knowledge are true and that all statements made on information and belief are believed to be true.

Albert S. Feng

Albert S. Feng, Inventor

January 20, 2005

Date

Chen Liu, Inventor

Date

Charissa R. Lansing

Charissa R. Lansing, Inventor

January 25, 2005

Date

William D. O'Brien Jr.

William D. O'Brien, Jr., Inventor

January 24, 2005

Date

Bruce C Wheeler

Bruce C. Wheeler, Inventor

January 24, 2005

Date

IN THE UNITED STATES PATENT OFFICE

In re patent application of:	)	
	)	Before the Examiner:
Albert S. Feng et al.	)	Andrew R. Graham
	)	
Application No. 09/193,058	)	
	)	Group Art Unit: 2644
Filed: November 16, 1998	)	
	)	
BINAURAL SIGNAL PROCESSING TECHNIQUES	)	
	)	

DECLARATION UNDER 37 CFR §1.131

We, the inventors of the above-indicated patent application, hereby declare as follows:

1. We have reviewed U.S. Patent Application No. 09/193,058 (the "'Subject Application"), which is attached in exhibit A as originally filed on November 16, 1998, and U.S. Patent Application No. 08/666,757, (the "Parent Application"), which is attached in exhibit B as filed June 19, 1996. The Parent Application issued on April 24, 2001 as U.S. Patent No. 6,222,927.

2. We have each reviewed claims 34, 35, 37-42, 44-48, 50-54, and 56-66 that have been or will be proposed for the Subject Application (the "Patent Claims"), a copy of which is attached as exhibit C.

3. We have compared the Patent Claims of exhibit C to the text of the Parent Application of exhibit B and the text of the Subject Application of Exhibit C. Based on this comparison, the Parent Application describes in writing the inventions defined by the Patent Claims in

the text common to the Subject Application. Specifically, the written description of the Parent Application appearing on pages 1-3 is substantially the same as the written description of the Subject Application appearing on page 1, line 13 – page 3, line 17; the written description of the Parent Application appearing on pages 8-17 is substantially the same as the written description of the Subject Application appearing on page 9 – page 17, line 27; and the written description of the Parent Application appearing on pages 18-19 (Experimental Section) is substantially the same as the written description of the Subject Application appearing on page 37 (Example One).

4. Research directed to extracting a desired acoustic signal from a noisy acoustic environment, such as that associated with the "cocktail party effect," began prior to September 1995 and continued at least into the year 1996 as evidenced by section 14 (pages 6-7) of the attached exhibit D. Exhibit D includes an Invention Disclosure Report prepared by the inventors, a cover letter dated May 9, 1996, and supplemental information for the Invention Disclosure Report, and has been partially redacted.

5. Based on information and belief, the cover letter dated May 9, 1996 forwarded the Invention Disclosure Report to the law firm that prepared and filed the Parent Application and the Subject Application. Such firm subsequently received the supplemental information of exhibit D before June 1996.

6. Based on information and belief, at least some of the records and/or files referenced in exhibit D (including that referenced in section 14(c)) cannot be located due to inadvertent

destruction or loss in connection with one or more equipment upgrades in years subsequent to the filing of the Parent Application.

7. The subtractive processing algorithm, accompanying mathematical formulae, and other aspects set forth in exhibit D were established at least as early as September 15, 1995. No later than this date, we formed in our minds a definite and permanent idea of the complete and operative inventions defined by the methods of the Patent Claims with the establishment of these concepts. We confirm our recollection of this timing by its chronological relationship to the information set forth in the documents of Exhibit E that all existed prior to September 1995.

8. From before September 18, 1995 through filing of the Parent Application on June 19, 1996, the inventors have diligently continued research, development, evaluation, and experimentation regarding the inventions defined by the Patent Claims. Such activity before September 18, 1995 is supported by at least the information set forth in exhibit D. Also, in exhibit D note the formation of one corresponding research team August of 1994 (section 14(a) followed by an initial written record September of 1995 (section 14 (b)). This initial record was prepared by inventor Chen Lui shortly after joining the research effort during late August of 1995 as corroborated by certain entries of exhibit E attached.

9. After September 18, 1995, activities continued with the preparation of a detailed research initiative proposal dated November 29, 1995, in which the subject matter of exhibit E corresponds to the text with the heading "aim #1." A copy of a draft of this proposal is

provided in exhibit F.

10. Before February 9, 1996, experimental activities included computer simulation of the processes described in at least independent claims 34, 46, 61, and 62 of the Patent Claims, as explained in section 7 (pages 4-5) of the Invention Disclosure Record of exhibit D. These efforts are also discussed in the communication to inventor, Dr. Albert Feng, on February 12, 1996 from Lynn Huerta, which is included as exhibit G. To those skilled in the art to which the inventions of the Patent Claims pertain, this type of simulation establishes performance of the corresponding inventions in the intended manner.

11. Before forwarding the Invention Disclosure Report of exhibit D to counsel on or about May 9, 1996, testing of an experimental prototype was conducted further establishing performance of the inventions of at least independent claims 34, 46, 61, and 62 of the Patent Claims in the intended manner. This experimentation is further detailed in section 10 of the Invention Disclosure Report of exhibit D. The experimental section of the Parent Application (pages 18-19) and the Subject Application (page 37) correspond to our experimental activities.

12. Based on information and belief, counsel that received the Invention Disclosure Report of exhibit D promptly reviewed it and arranged an interview by telephone with at least one of the inventors, Dr. Albert Feng, to further discuss the information contained therein. The supplement to the Invention Disclosure Report of exhibit D was subsequently

received by such counsel before June 1996.

13. Based on information and belief, counsel reviewed the Invention Disclosure Report and prepared the Parent Application from the materials provided between May 9, 1996 and its filing date of June 19, 1996.

14. The undersigned, being hereby warned that willful false statements and the like are punishable by fine or imprisonment, or both (18 U. S C. 1001), and may jeopardize the validity of the application or any patent issuing thereon, declares that all statements made of his/her own knowledge are true and that all statements made on information and belief are believed to be true.

____
Albert S. Feng, Inventor

____
Date

Chen Liu
Chen Liu, Inventor

Jan. 24, 2005
Date

Charissa R. Lansing, Inventor

Date

William D. O'Brien, Jr., Inventor

Date

Bruce C. Wheeler, Inventor

Date

22010-127/16-13860

COPY

BINAURAL SIGNAL PROCESSING TECHNIQUES

CROSS-REFERENCE TO RELATED APPLICATION

This application is a continuation-in-part of pending United States Patent Application Serial No. 08/666,757, filed on June 19, 1996 by the same inventive entity, and entitled BINAURAL SIGNAL PROCESSING SYSTEM AND METHOD.

BACKGROUND OF THE INVENTION

The present invention is directed to the processing of acoustic signals, and more particularly, but not exclusively, relates to the localization and extraction of acoustic signals emanating from different sources.

The difficulty of extracting a desired signal in the presence of interfering signals is a long-standing problem confronted by acoustic engineers. This problem impacts the design and construction of many kinds of devices such as systems for voice recognition and intelligence gathering. Especially troublesome is the separation of desired sound from unwanted sound with hearing aid devices. Generally, hearing aid devices do not permit selective amplification of a desired sound when contaminated by noise from a nearby source -- particularly when the noise is more intense. This problem is even more severe when the desired sound is a speech signal and the nearby noise is also a speech signal produced by multiple talkers (e.g. babble). As used herein, "noise" refers not only to random or nondeterministic signals, but also to undesired signals and signals interfering with the perception of a desired signal.

One attempted solution to this problem has been the application of a single, highly directional microphone to enhance directionality of the hearing aid receiver. This approach has only a very limited capability. As a result, spectral subtraction, comb filtering, and speech-production modeling have been explored to enhance single microphone performance. Nonetheless, these approaches still generally fail to improve intelligibility of a desired speech signal, particularly when the signal and noise sources are in close proximity.

Another approach has been to arrange a number of microphones in a selected spatial relationship to form a type of directional detection beam. Unfortunately, when limited to a size practical for hearing aids, beam forming arrays also have limited capacity to separate signals that are close together -- especially if the noise is more intense than the desired speech signal. In addition, in the case of one noise source in a less reverberant environment, the noise cancellation provided by the beam-former varies with the location of the noise source in relation to the microphone array. R.W. Stadler and W.M. Rabinowitz, On the Potential of Fixed Arrays for Hearing Aids, 94 Journal Acoustical Society of America 1332 (September 1993), and W. Soede et al., Development of a Directional Hearing Instrument Based on Array Technology, 94 Journal of Acoustical Society of America 785 (August 1993) are cited as additional background concerning the beam forming approach.

Still another approach has been the application of two microphones displaced from one another to provide two signals to emulate certain aspects of the binaural hearing system common to humans and many types of animals. Although certain aspects of biologic binaural hearing are not fully understood, it is believed that the ability to localize sound sources is based on evaluation by the auditory system of binaural time delays and sound levels across different frequency bands associated with each of the two sound signals. The localization of sound sources with systems based on these interaural time and intensity differences is discussed in W. Lindemann, Extension of a Binaural Cross-Correlation Model by Contralateral Inhibition - I. Simulation of Lateralization for Stationary Signals, 80 Journal of the Acoustical Society of America 1608 (December 1986).

The localization of multiple acoustic sources based on input from two microphones presents several significant challenges,.as does the separation of a desired signal once the sound sources are localized. For example, the system set forth in Markus Bodden, Modeling Human Sound-Source Localization and the Cocktail-Party-Effect, 1 Acta Acustica 43 (February/April 1993) employs a Wiener filter including a windowing process in an attempt to derive a desired signal from binaural input signals once the location of the desired signal has been established. Unfortunately, this approach results in significant deterioration of desired speech fidelity. Also, the system has only been demonstrated to suppress noise of equal intensity to the desired signal at an azimuthal separation of at least 30 degrees. A more intense noise emanating from a source spaced closer than 30 degrees from the desired source continues to present a problem. Moreover, the proposed algorithm of the Bodden system is computationally intense -- posing a serious question of whether it can be practically embodied in a hearing aid device.

Another example of a two microphone system is found in D. Banks, Localisation and Separation of Simultaneous Voices with Two Microphones, IEE Proceedings-I, 140 (1993). This system employs a windowing technique to estimate the location of a sound source when there are nonoverlapping gaps in its spectrum compared to the spectrum of interfering noise. This system cannot perform localization when wide-band signals lacking such gaps are involved. In addition, the Banks article fails to provide details of the algorithm for reconstructing the desired signal. U.S. Patent Nos. 5,479,522 to Lindemann et al.; 5,325,436 to Soli et al.; 5,289,544 to Franklin; and 4,773,095 to Zwicker et al. are cited as sources of additional background concerning dual microphone hearing aid systems.

Effective localization is also often hampered by ambiguous positional information that results above certain frequencies related to the spacing of the input microphones. This problem was recognized in Stern, R. M., Zeiberg, A. S., and Trahiotis, C. "Lateralization of complex binaural stimuli: A weighted-image model," J. Acoust. Soc. Am. 84, 156-165 (1988).

Thus, a need remains for more effective localization and extraction techniques -- especially for use with binaural systems. The present invention meets these needs and offers other significant benefits and advantages.

SUMMARY OF THE INVENTION

The present invention relates to the processing of acoustic signals. Various aspects of the invention are novel, nonobvious, and provide various advantages. While the actual nature of the invention covered herein can only be determined with reference to the claims appended hereto, selected forms and features of the preferred embodiments as disclosed herein are described briefly as follows.

One form of the present invention includes a signal processing technique for localizing and characterizing each of a number of differently located acoustic sources. Detection of the sources is performed with two sensors that are spaced apart. Each, or one particular selected source may be extracted, while suppressing the output of the other sources. A variety of applications may benefit from this technique including hearing aids, sound location mapping or tracking devices, and voice recognition equipment, to name a few.

In another form, a first signal is provided from a first acoustic sensor and a second signal from a second acoustic sensor spaced apart from the first acoustic sensor. The first and second signals each correspond to a composite of two or more acoustic sources that, in turn, include a plurality of interfering sources and a desired source. The interfering sources are localized by processing of the first and second signals to provide a corresponding number of interfering source signals. These signals each include a number of frequency components. One or more the frequency components are suppressed for each of the interfering source signals. This approach facilitates nulling a different frequency component for each of a number of noise sources with two input sensors.

A further form of the present invention is a processing system having a pair of sensors and a delay operator responsive to a pair of input signals from the sensors to generate a number of delayed signals therefrom. The system also has a localization operator responsive to the delayed signals to localize the interfering sources relative to the location of the sensors and provide a plurality of interfering source signals each represented by a number of frequency components. The system further includes an extraction operator that serves to suppress selected frequency components for each of the interfering source signals and extract a desired signal corresponding to a desired source. An output device responsive to the desired signal is also included that provides an output representative of the desired

source. This system may be incorporated into a signal processor coupled to the sensors to facilitate localizing and suppressing multiple noise sources when extracting a desired signal.

Still another form is responsive to position-plus-frequency attributes of sound sources. It includes positioning a first acoustic sensor and a second acoustic sensor to detect a plurality of differently located acoustic sources. First and second signals are generated by the first and second sensors, respectively, that receive stimuli from the acoustic sources. A number of delayed signal pairs are provided from the first and second signals that each correspond to one of a number of positions relative to the first and second sensors. The sources are localized as a function of the delayed signal pairs and a number of coincidence patterns. These patterns are position and frequency specific, and may be utilized to recognize and correspondingly accumulate position data estimates that map to each true source position. As a result, these patterns may operate as filters to provide better localization resolution and eliminate spurious data.

In yet another form, a system includes two sensors each configured to generate a corresponding first or second input signal and a delay operator responsive to these signals to generate a number of delayed signals each corresponding to one of a number of positions relative to the sensors. The system also includes a localization operator responsive to the delayed signals for determining the number of sound source localization signals. These localization signals are determined from the delayed signals and a number of coincidence patterns that each correspond to one of the positions. The patterns each relate frequency varying sound source location information caused by ambiguous phase multiples to a corresponding position to improve acoustic source localization. The system also has an output device responsive to the localization signals to provide an output corresponding to at least one of the sources.

A further form utilizes two sensors to provide corresponding binaural signals from which the relative separation of a first acoustic source from a second acoustic source may be established as a function of time, and the spectral content of a desired acoustic signal from the first source may be representatively extracted. Localization and identification of the spectral content of the desired acoustic signal may be performed concurrently. This form may also successfully extract the desired acoustic signal even if a nearby noise source is of greater relative intensity.

Another form of the present invention employs a first and second sensor at different locations to provide a binaural representation of an acoustic signal which includes a desired signal emanating from a selected source and interfering signals emanating from several

interfering sources. A processor generates a discrete first spectral signal and a discrete second spectral signal from the sensor signals. The processor delays the first and second spectral signals by a number of time intervals to generate a number of delayed first signals and a number of delayed second signals and provide a time increment signal. The time increment signal corresponds to separation of the selected source from the noise source. The processor generates an output signal as a function of the time increment signal, and an output device responds to the output signal to provide an output representative of the desired signal.

An additional form includes positioning a first and second sensor relative to a first signal source with the first and second sensor being spaced apart from each other and a second signal source being spaced apart from the first signal source. A first signal is provided from the first sensor and a second signal is provided from the second sensor. The first and second signals each represents a composite acoustic signal including a desired signal from the first signal source and unwanted signals from other sound sources. A number of spectral signals are established from the first and second signals as functions of a number of frequencies. A member of the spectral signals representative of position of the second signal source is determined, and an output signal is generated from the member which is representative of the first signal source. This feature facilitates extraction of a desired signal from a spectral signal determined as part of the localization of the interfering source. This approach avoids the extensive post-localization computations required by many binaural systems to extract a desired signal.

Accordingly, it is one object of the present invention to provide for the enhanced localization of multiple acoustic sources.

It is another object to extract a desired acoustic signal from a noisy environment caused by a number of interfering sources.

An additional object is to provide a system for the localization and extraction of acoustic signals by detecting a combination of these signals with two differently located sensors.

Further objects, features, aspects, benefits, forms, and advantages of the present invention shall become apparent from the detailed drawings and descriptions provided herein.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagrammatic view of a system of one embodiment of the present invention.

FIG. 2 is a signal flow diagram further depicting selected aspects of the system of FIG. 1.

FIG. 3 is schematic representation of the dual delay line of FIG. 2.

FIGS. 4A and 4B depict other embodiments of the present invention corresponding to hearing aid and computer voice recognition applications, respectively.

FIG. 5 is a graph of a speech signal in the form of a sentence about 2 seconds long.

FIG. 6 is a graph of a composite signal including babble noise and the speech signal of FIG. 5 at a 0 dB signal-to-noise ratio with the babble noise source at about a 60 azimuth relative to the speech signal source.

FIG. 7 is a graph of a signal representative of the speech signal of FIG. 5 after extraction from the composite signal of FIG. 6.

FIG. 8 is a graph of a composite signal including babble noise and the speech signal of FIG. 5 at a -30 dB signal-to-noise ratio with the babble noise source at a 2 degree azimuth relative to the speech signal source.

FIG. 9 is a graphic depiction of a signal representative of the sample speech signal of FIG. 5 after extraction from the composite signal of FIG. 8.

FIG. 10 is a signal flow diagram of another embodiment of the present invention.

FIG. 11 is a partial, signal flow diagram illustrating selected aspects of the dual delay lines of FIG. 10 in greater detail.

FIG. 12 is a diagram illustrating selected geometric features of the embodiment illustrated in FIG. 10 for a representative example of one of a number of sound sources.

FIG. 13 is a signal flow diagram illustrating selected aspects of the localization operator of FIG. 10 in greater detail.

FIG. 14 is a diagram illustrating yet another embodiment of the present invention.

FIG. 15 is a signal flow diagram further illustrating selected aspects of the embodiment of FIG. 14.

FIG. 16 is a signal flow diagram illustrating selected aspects of the localization operator of FIG. 15 in greater detail.

FIG. 17 is a graph illustrating a plot of coincidence loci for two sources.

FIG. 18 is a graph illustrating coincidence patterns for azimuth positions corresponding to -75°, 0°, 20°, and 75°.

FIGs. 19-22 are tables depicting experimental results obtained with the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENT

For the purposes of promoting an understanding of the principles of the invention, reference will now be made to the embodiment illustrated in the drawings and specific language will be used to describe the same. It will nevertheless be understood that no limitation of the scope of the invention is thereby intended. Any alterations and further modifications in the described device, and any further applications of the principles of the invention as described herein are contemplated as would normally occur to one skilled in the art to which the invention relates.

Fig. 1 illustrates an acoustic signal processing system 10 of one embodiment of the present invention. System 10 is configured to extract a desired acoustic signal from source 12 despite interference or noise emanating from nearby source 14. System 10 includes a pair of acoustic sensors 22, 24 configured to detect acoustic excitation that includes signals from sources 12, 14. Sensors 22, 24 are operatively coupled to processor 30 to process signals received therefrom. Also, processor 30 is operatively coupled to output device 90 to provide a signal representative of a desired signal from source 12 with reduced interference from source 14 as compared to composite acoustic signals presented to sensors 22, 24 from sources 12, 14.

Sensors 22, 24 are spaced apart from one another by distance D along lateral axis T. Midpoint M represents the half way point along distance D from sensor 22 to sensor 24. Reference axis R1 is aligned with source 12 and intersects axis T perpendicularly through midpoint M. Axis N is aligned with source 14 and also intersects midpoint M. Axis N is positioned to form angle A with reference axis R1. Fig. 1 depicts an angle A of about 20 degrees. Notably, reference axis R1 may be selected to define a reference azimuthal position of zero degrees in an azimuthal plane intersecting sources 12, 14; sensors 22, 24; and containing axes T, N, R1. As a result, source 12 is "on-axis" and source 14, as aligned with axis N, is "off-axis." Source 14 is illustrated at about a 20 degree azimuth relative to source 12.

Preferably sensors 22, 24 are fixed relative to each other and configured to move in tandem to selectively position reference axis R1 relative to a desired acoustic signal source. It is also preferred that sensors 22, 24 be microphones of a conventional variety, such as omnidirectional dynamic microphones. In other embodiments, a different sensor type may be utilized as would occur to one skilled in the art.

Referring additionally to FIG. 2, a signal flow diagram illustrates various processing stages for the embodiment shown in FIG. 1. Sensors 22, 24 provide analog signals Lp(t) and Rp(t) corresponding to the left sensor 22, and right sensor 24, respectively. Signals Lp(t) and Rp(t) are initially input to processor 30 in separate processing channels L and R. For each channel L, R, signals Lp(t) and Rp(t) are conditioned and filtered in stages 32a, 32b to reduce aliasing, respectively. After filter stages 32a, 32b, the conditioned signals Lp(t), Rp(t) are input to corresponding Analog to Digital (A/D) converters 34a, 34b to provide discrete signals Lp(k), Rp(k), where k indexes discrete sampling events. In one embodiment, A/D stages 34a, 34b sample signals Lp(t) and Rp(t) at a rate of at least twice the frequency of the upper end of the audio frequency range to assure a high fidelity representation of the input signals.

Discrete signals Lp(k) and Rp(k) are transformed from the time domain to the frequency domain by a short-term Discrete Fourier Transform (DFT) algorithm in stages 36a, 36b to provide complex-valued signals XLp(m) and XRp(m). Signals XLp(m) and XRp(m) are evaluated in stages 36a, 36b at discrete frequencies f_m, where m is an index (m=1 to m=M) to discrete frequencies, and index p denotes the short-term spectral analysis time frame. Index p is arranged in reverse chronological order with the most recent time frame being p =1, the next most recent time frame being p = 2, and so forth. Preferably, frequencies M encompass the audible frequency range and the number of samples employed in the short-term analysis is selected to strike an optimum balance between processing speed limitations and desired resolution of resulting output signals. In one embodiment, an audio range of 0.1 to 6 kHz is sampled in A/D stages 34a, 34b at a rate of at least 12.5 kHz with 512 samples per short-term spectral analysis time frame. In alternative embodiments, the frequency domain analysis may be provided by an analog filter bank employed before A/D stages 34a, 34b. It should be understood that the spectral signals XLp(m) and XRp(m) may be represented as arrays each having a 1xM dimension corresponding to the different frequencies f_m.

Spectral signals XLp(m) and XRp(m) are input to dual delay line 40 as further detailed in FIG. 3. FIG. 3 depicts two delay lines 42, 44 each having N number of delay stages. Each delay line 42, 44 is sequentially configured with delay stages D_1 through D_N. Delay lines 42, 44 are configured to delay corresponding

input signals in opposing directions from one delay stage to the next, and generally correspond to the dual hearing channels associated with a natural binaural hearing process. Delay stages D_1, D_2, D_3, . . ., D_{N-2}, D_{N-1}, and D_N each delay an input signal by corresponding time delay increments τ_1, τ_2, τ_3, . . ., τ_{N-2}, τ_{N-1}, and τ_N, (collectively designated τ_i), where index i goes from left to right. For delay line 42, XLp(m) is alternatively designated $XLp^1(m)$. $XLp^1(m)$ is sequentially delayed by time delay increments τ_1, τ_2, τ_3, . . ., τ_{N-2}, τ_{N-1}, and τ_N to produce delayed outputs at the taps of delay line 42 which are respectively designated $XLp^2(m)$, $XLp^3(m)$, $Xlp^4(m)$, . . ., $XLp^{N-1}(m)$, $XLp^N(m)$, and $XLp^{N+1}(m)$; and collectively designated $XLp^i(m)$). For delay line 44, XRp(m) is alternatively designated $XRp^{N+1}(m)$. $XRp^{N+1}(m)$ is sequentially delayed by time delay increments τ_1, τ_2, τ_3, . . ., τ_{N-2}, τ_{N-1}, and τ_N to produce delayed outputs at the taps of delay line 44 which are respectively designated: $XRp^N(m)$, $XRp^{N-1}(m)$, $XRp^{N-2}(m)$, . . ., $XLp^3(m)$, $XLp^2(m)$, and $Xlp^1(m)$; and collectively designated $XRp^i(m)$. The input spectral signals and the signals from delay line 42, 44 taps are arranged as input pairs to operation array 46. A pair of taps from delay lines 42, 44 is illustrated as input pair P in FIG. 3.

Operation array 46 has operation units (OP) numbered from 1 to N+1, depicted as OP1, OP2, OP3, OP4,..., OPN-2, OPN-1, OPN, OPN+1 and collectively designated operations OPi. Input pairs from delay lines 42, 44 correspond to the operations of array 46 as follows: OP1[$XLp^1(m)$, $XRp^1(m)$], OP2[$XLp^2(m)$, $XRp^2(m)$], OP3[$XLp^3(m)$, $XRp^3(m)$], OP4[$XLp^4(m)$, $XRp^4(m)$],..., OPN-2[$XLp^{(N-2)}(m)$, $XRp^{(N-2)}(m)$], OPN-1[$XLp^{(N-1)}(m)$, $XRp^{(N-1)}(m)$], OPN[$XLp^N(m)$, $XRp^N(m)$], and OPN+1[$XLp^{(N+1)}(m)$, $XRp^{(N+1)}(m)$]; where OPi[$XLp^i(m)$, $XRp^i(m)$] indicates that OPi is determined as a function of input pair $XLp^i(m)$, $XRp^i(m)$. Correspondingly, the outputs of operation array 46 are $Xp^1(m)$, $Xp^2(m)$, $Xp^3(m)$, $Xp^4(m)$, ..., $Xp^{(N-2)}(m)$, $Xp^{(N-1)}(m)$, $Xp^N(m)$, and $Xp^{(N+1)}(m)$ (collectively designated $Xp^i(m)$).

For $i = 1$ to $i \leq N/2$, operations for each OPi of array 46 are determined in accordance with complex expression 1 (CE1) as follows:

$$Xp^i(m) = \frac{XLp^i(m) - XRp^i(m)}{\exp[-j2\pi(\tau_i+...+\tau_{N/2})f_m] - \exp[j2\pi(\tau_{((N/2)+1)}+...+\tau_{(N-i+1)})f_m]},$$

where exp[argument] represents a natural exponent to the power of the argument, and imaginary number j is the square root of -1. For $i > ((N/2)+1)$ to $i = N+1$, operations of operation array 46 are determined in accordance complex expression 2 (CE2) as follows:

$$Xp^i(m) = \frac{XLp^i(m) - XRp^i(m)}{\exp[j2\pi(\tau_{((N/2)+1)}+...+\tau_{(i-1)})f_m] - \exp[-j2\pi(\tau_{(N-i+2)}+...+\tau_{N/2})f_m]},$$

where exp[argument] represents a natural exponent to the power of the argument, and imaginary number j is the square root of -1. For $i = (N/2)+1$, neither CE1 nor CE2 is performed.

An example of the determination of the operations for $N = 4$ ($i=1$ to $i=N+1$) is as follows:

$i = 1$, CE1 applies as follows:

$$Xp^1(m) = \frac{XLp^1(m) - XRp^1(m)}{\exp[-j2\pi(\tau_1+\tau_2)f_m] - \exp[j2\pi(\tau_3+\tau_4)f_m]};$$

$i = 2 \leq (N/2)$, CE1 applies as follows:

$$Xp^2(m) = \frac{XLp^2(m) - XRp^2(m)}{\exp[-j2\pi(\tau_2)f_m] - \exp[j2\pi(\tau_3)f_m]};$$

$i = 3$: Not applicable, $(N/2) < i \leq ((N/2)+1)$;

$i = 4$, CE2 applies as follows:

$$Xp^4(m) = \frac{XLp^4(m) - XRp^4(m)}{\exp[j2\pi(\tau_3)f_m] - \exp[-j2\pi(\tau_2)f_m]}; \text{ and,}$$

$i = 5$, CE2 applies as follows:

$$Xp^5(m) = \frac{XLp^5(m) - XRp^5(m)}{\exp[j2\pi(\tau_3+\tau_4)f_m] - \exp[-j2\pi(\tau_1+\tau_2)f_m]}.$$

Referring to FIGS. 1-3, each OPi of operation array 46 is defined to be representative of a different azimuthal position relative to reference axis R. The "center" operation, OPi where $i = ((N/2)+1)$, represents the location of the reference axis and source 12. For the example N=4, this center operation corresponds to $i = 3$. This arrangement is analogous to the different interaural time differences associated with a natural binaural hearing system. In these natural systems, there is a relative position in each sound passageway within the ear that corresponds to a maximum "in phase" peak for a given sound source. Accordingly, each operation of array 46 represents a position corresponding to a potential azimuthal or angular position range for a sound source, with the center operation representing a source at the zero azimuth -- a source aligned with reference axis R. For an environment having a single source without noise or interference, determining the signal pair with the maximum strength may be sufficient to locate the source with little additional processing; however, in noisy or multiple source environments, further processing may be needed to properly estimate locations.

It should be understood that dual delay line 40 provides a two dimensional matrix of outputs with N+1 columns corresponding to $Xp^i(m)$, and M rows corresponding to each discrete frequency f_m of $Xp^i(m)$. This (N+1)xM matrix is determined for each

short-term spectral analysis interval p. Furthermore, by subtracting $XRp^{i}(m)$ from $XLp^{i}(m)$, the denominator of each expression CE1, CE2 is arranged to provide a minimum value of $Xp^{i}(m)$ when the signal pair is "in-phase" at the given frequency f_m. Localization stage 70 uses this aspect of expressions CE1, CE2 to evaluate the location of source 14 relative to source 12.

Localization stage 70 accumulates P number of these matrices to determine the $Xp^{i}(m)$ representative of the position of source 14. For each column i, localization stage 70 performs a summation of the amplitude of $|Xp^{i}(m)|$ to the second power over frequencies f_m from m=1 to m=M. The summation is then multiplied by the inverse of M to find an average spectral energy as follows:

$$Xavgp^{i} = (1/M) \sum_{m=1}^{M} |Xp^{i}(m)|^{2}.$$

The resulting averages, $Xavgp^{i}$ are then time averaged over the P most recent spectral-analysis time frames indexed by p in accordance with:

$$X^{i} = \sum_{p=1}^{P} \gamma p \, Xavgp^{i} ,$$

where γp are empirically determined weighting factors. In one embodiment, the γp factors are preferably between 0.85^{p} and 0.90^{p}, where p is the short-term spectral analysis time frame index. The X^{i} are analyzed to determine the minimum value, $\min(X^{i})$. The index i of $\min(X^{i})$, designated "I," estimates the column representing the azimuthal location of source 14 relative to source 12.

It has been discovered that the spectral content of a desired signal from source 12, when approximately aligned with reference axis R1, can be estimated from $Xp^{I}(m)$. In other words, the spectral signal output by array 46 which most closely corresponds to the relative location of the "off-axis" source 14 contemporaneously provides a spectral representation of a signal emanating from source 12. As a result, the signal processing of dual delay line 40 not only facilitates localization of source

14, but also provides a spectral estimate of the desired signal with only minimal post-localization processing to produce a representative output.

Post-localization processing includes provision of a designation signal by localization stage 70 to conceptual "switch" 80 to select the output column $Xp^I(m)$ of the dual delay line 40. The $Xp^I(m)$ is routed by switch 80 to an inverse Discrete Fourier Transform algorithm (Inverse DFT) in stage 82 for conversion from a frequency domain signal representation to a discrete time domain signal representation denoted as s(k). The signal estimate s(k) is then converted by Digital to Analog (D/A) converter 84 to provide an output signal to output device 90.

Output device 90 amplifies the output signal from processor 30 with amplifier 92 and supplies the amplified signal to speaker 94 to provide the extracted signal from a source 12.

It has been found that interference from off-axis sources separated by as little as 2 degrees from the on axis source may be reduced or eliminated with the present invention -- even when the desired signal includes speech and the interference includes babble. Moreover, the present invention provides for the extraction of desired signals even when the interfering or noise signal is of equal or greater relative intensity. By moving sensors 22, 24 in tandem the signal selected to be extracted may correspondingly be changed. Moreover, the present invention may be employed in an environment having many sound sources in addition to sources 12, 14. In one alternative embodiment, the localization algorithm is configured to dynamically respond to relative positioning as well as relative strength, using automated learning techniques. In other embodiments, the present invention is adapted for use with highly directional microphones, more than two sensors to simultaneously extract multiple signals, and various adaptive amplification and filtering techniques known to those skilled in the art.

The present invention greatly improves computational efficiency compared to conventional systems by determining a spectral signal representative of the desired signal as part of the localization processing. As a result, an output signal characteristic of a desired signal from source 12 is determined as a function of the signal pair $XLp^I(m)$, $XRp^I(m)$ corresponding to the separation of source 14 from source 12. Also, the exponents in the denominator of CE1, CE2 correspond to phase

difference of frequencies f_m resulting from the separation of source 12 from 14. Referring to the example of N=4 and assuming that I=1, this phase difference is $-2\pi(\tau_1+\tau_2)/f_m$ (for delay line 42) and $2\pi(\tau_3+\tau_4)/f_m$ (for delay line 44) and corresponds to the separation of the representative location of off-axis source 14 from the on-axis source 12 at i=3. Likewise the time increments, $\tau_1+\tau_2$ and $\tau_3+\tau_4$, correspond to the separation of source 14 from source 12 for this example. Thus, processor 30 implements dual delay line 40 and corresponding operational relationships CE1, CE2 to provide a means for generating a desired signal by locating the position of an interfering signal source relative to the source of the desired signal.

It is preferred that τ_i be selected to provide generally equal azimuthal positions relative to reference axis R. In one embodiment, this arrangement corresponds to the values of τ_i changing about 20% from the smallest to the largest value. In other embodiments, τ_i are all generally equal to one another, simplifying the operations of array 46. Notably, the pair of time increments in the numerator of CE1, CE2 corresponding to the separation of the sources 12 and 14 become approximately equal when all values τ_i are generally the same.

Processor 30 may be comprised of one or more components or pieces of equipment. The processor may include digital circuits, analog circuits, or a combination of these circuit types. Processor 30 may be programmable, an integrated state machine, or utilize a combination of these techniques. Preferably, processor 30 is a solid state integrated digital signal processor circuit customized to perform the process of the present invention with a minimum of external components and connections. Similarly, the extraction process of the present invention may be performed on variously arranged processing equipment configured to provide the corresponding functionality with one or more hardware modules, firmware modules, software modules, or a combination thereof. Moreover, as used herein, "signal" includes, but is not limited to, software, firmware, hardware, programming variable, communication channel, and memory location representations.

Referring to FIG. 4A, one application of the present invention is depicted as hearing aid system 110. System 110 includes eyeglasses G with microphones 122 and 124 fixed to glasses G and displaced from one another. Microphones 122, 124 are

operatively coupled to hearing aid processor 130. Processor 130 is operatively coupled to output device 190. Output device 190 is positioned in ear E to provide an audio signal to the wearer.

Microphones 122, 124 are utilized in a manner similar to sensors 22, 24 of the embodiment depicted by FIGS 1-3. Similarly, processor 130 is configured with the signal extraction process depicted in of FIGS. 1-3. Processor 130 provides the extracted signal to output device 190 to provide an audio output to the wearer. The wearer of system 110 may position glasses G to align with a desired sound source, such as a speech signal, to reduce interference from a nearby noise source off axis from the midpoint between microphones 122, 124. Moreover, the wearer may select a different signal by realigning with another desired sound source to reduce interference from a noisy environment.

Processor 130 and output device 190 may be separate units (as depicted) or included in a common unit worn in the ear. The coupling between processor 130 and output device 190 may be an electrical cable or a wireless transmission. In one alternative embodiment, sensors 122, 124 and processor 130 are remotely located and are configured to broadcast to one or more output devices 190 situated in the ear E via a radio frequency transmission or other conventional telecommunication method.

FIG. 4B shows a voice recognition system 210 employing the present invention as a front end speech enhancement device. System 210 includes personal computer C with two microphones 222, 224 spaced apart from each other in a predetermined relationship. Microphones 222, 224 are operatively coupled to a processor 230 within computer C. Processor 230 provides an output signal for internal use or responsive reply via speakers 294a, 294b or visual display 296. An operator aligns in a predetermined relationship with microphones 222, 224 of computer C to deliver voice commands. Computer C is configured to receive these voice commands, extracting the desired voice command from a noisy environment in accordance with the process system of FIGS. 1-3.

Referring to Figs. 10-13, signal processing system 310 of another embodiment of the present invention is illustrated. Reference numerals of system 310 that are the same as those of system 10 refer to like features. The signal flow diagram of FIG. 10 corresponds to various signal processing techniques of system 310. Fig. 10 depicts left "L" and right "R" input channels for signal processor 330 of system 310. Channels L, R each include an acoustic sensor 22, 24 that provides an input signal $x_{Ln}(t)$, $x_{Rn}(t)$, respectively. Input signals $x_{Ln}(t)$ and $x_{Rn}(t)$ correspond to composites of sounds from multiple acoustic sources located

within the detection range of sensors 22, 24. As described in connection with FIG. 1 of system 10, it is preferred that sensors 22, 24 be standard microphones spaced apart from each other at a predetermined distance D. In other embodiments a different sensor type or arrangement may be employed as would occur to those skilled in the art.

Sensors 22, 24 are operatively coupled to processor 330 of system 310 to provide input signals $x_{Ln}(t)$ and $x_{Rn}(t)$ to A/D converters 34a, 34b. A/D converters 34a, 34b of processor 330 convert input signals $x_{Ln}(t)$ and $x_{Rn}(t)$ from an analog form to a discrete form as represented as $x_{Ln}(k)$ and $x_{Rn}(k)$, respectively; where "t" is the familiar continuous time domain variable and "k" is the familiar discrete sample index variable. A corresponding pair of preconditioning filters (not shown) may also be included in processor 330 as described in connection with system 10.

Digital Fourier Transform (DFT) stages 36a, 36b receive the digitized input signal pair $x_{Ln}(k)$ and $x_{Rn}(k)$ from converters 34a, 34b, respectively. Stages 36a, 36b transform input signals as $x_{Ln}(k)$ and $x_{Rn}(k)$ into spectral signals designated $X_{Ln}(m)$ and $X_{Rn}(m)$ using a short term discrete Fourier transform algorithm. Spectral signals $X_{Ln}(m)$ and $X_{Rn}(m)$ are expressed in terms of a number of discrete frequency components indexed by integer m; where m=1, 2, ..., M. Also, as used herein, the subscripts L and R denote the left and right channels, respectively, and n indexes time frames for the discrete Fourier transform analysis.

Delay operator 340 receives spectral signals $X_{Ln}(m)$ and $X_{Rn}(m)$ from stages 36a, 36b, respectively. Delay operator 340 includes a number of dual delay lines (DDLs) 342 each corresponding to a different one of the component frequencies indexed by m. Thus, there are M different dual delay lines 342 utilized. However, only dual delay lines 342 corresponding to m=1 and $m=M$ are shown in Fig. 10 to preserve clarity. The remaining dual delay lines corresponding to m=2 through $m=(M-1)$ are represented by an ellipsis to preserve clarity. Alternatively, delay operator 340 may be described as a single dual delay line that simultaneously operates on M frequencies like dual delay line 40 of system 10.

The pair of frequency components from DFT stages 36a, 36b corresponding to a given value of m are inputs into a corresponding one of dual delay lines 342. For the examples illustrated in Fig. 10, spectral signal component pair $X_{Ln}(m=1)$ and $X_{Rn}(m=1)$ is sent to the upper dual delay line 342 for the frequency corresponding to m=1; and spectral signal component pair $X_{Ln}(m=M)$ and $X_{Rn}(m=M)$ is sent to the lower dual delay line 342 for the frequency corresponding to $m=M$. Likewise, common frequency component pairs of

$X_{Ln}(m)$ and $X_{Rn}(m)$ for frequencies corresponding to m=2 through m=(M-1) are each sent to a corresponding dual delay line as represented by ellipses to preserve clarity.

Referring additionally to Fig. 11, certain features of dual delay line 342 are further illustrated. Each dual delay line 342 includes a left channel delay line 342a receiving a corresponding frequency component input from DFT stage 36a and right channel delay line 342b receiving a corresponding frequency component input from DFT stage 36b. Delay lines 342a, 342b each include an odd number I of delay stages 344 indexed by i=1, 2, ..., I. The I number of delayed signal pairs are provided on outputs 345 of delay stages 344 and are correspondingly sent to complex multipliers 346. There is one multiplier 346 corresponding to each delay stage 344 for each delay line 342a, 342b. Multipliers 346 provide equalization weighting for the corresponding outputs of delay stages 344. Each delayed signal pair from corresponding outputs 345 has one member from a delay stage 344 of left delay line 342a and the other member from a delay stage 344 of right delay line 342b. Complex multipliers 346 of each dual delay line 342 output corresponding products of the I number of delayed signal pairs along taps 347. The I number of signal pairs from taps 347 for each dual delay line 342 of operator 340 are input to signal operator 350.

For each dual delay line 342, the I number of pairs of multiplier taps 347 are each input to a different Operation Array (OA) 352 of operator 350. Each pair of taps 347 is provided to a different operation stage 354 within a corresponding operation array 352. In Fig. 11, only a portion of delay stages 344, multipliers 346, and operation stages 354 are shown corresponding to the two stages at either end of delay lines 342a, 342b and the middle stages of delay lines 342a, 342b. The intervening stages follow the pattern of the illustrated stages and are represented by ellipses to preserve clarity.

For an arbitrary frequency ω_m, delay times τ_i are given by equation (1) as follows:

$$\tau_i = \frac{\mathrm{ITD}_{max}}{2}\sin\left(\frac{i-1}{I-1}\pi - \frac{\pi}{2}\right), \qquad i=1, \ldots, I \qquad (1)$$

where, i is the integer delay stage index in the range (i=1, ..., I); ITD_{max} = D/c is the maximum Intermicrophone Time Difference; D is the distance between sensors 22, 24; and c is the speed of sound. Further, delay times τ_i are antisymmetric with respect to the midpoint of the delay stages corresponding to $i=(I+1)/2$ as indicated in the following equation (2):

$$\tau_{I-i+1} = \frac{\mathrm{ITD}_{max}}{2}\sin\left[\frac{(I-i+1)-1}{I-1}\pi - \frac{\pi}{2}\right] = -\frac{\mathrm{ITD}_{max}}{2}\sin\left(\frac{i-1}{I-1}\pi - \frac{\pi}{2}\right) = -\tau_i \qquad (2)$$

The azimuthal plane may be uniformly divided into I sectors with the azimuth position of each resulting sector being given by equation (3) as follows:

$$\theta_i = \frac{i-1}{I-1}180^\circ - 90^\circ, \qquad i=1, \ldots, I. \tag{3}$$

The azimuth positions in auditory space may be mapped to corresponding delayed signal pairs along each dual delay line 342 in accordance with equation (4) as follows:

$$\tau_i = \frac{\mathrm{ITD}_{max}}{2}\sin\theta_i, \qquad i=1, \ldots, I. \tag{4}$$

The dual delay-line structure is similar to the embodiment of system 10, except that a different dual delay line is represented for each value of m and multipliers 346 have been included to multiply each corresponding delay stage 344 by an appropriate one of equalization factors $\alpha_i(m)$; where i is the delay stage index previously described. Preferably, elements $\alpha_i(m)$ are selected to compensate for differences in the noise intensity at sensors 22, 24 as a function of both azimuth and frequency.

One preferred embodiment for determining equalization factors $\alpha_i(m)$ assumes amplitude compensation is independent of frequency, regarding any departure from this model as being negligible. For this embodiment, the amplitude of the received sound pressure $|\mathbf{p}|$ varies with the source-receiver distance r in accordance with equations (A1) and (A2) as follows:

$$|\mathbf{p}| \propto \frac{1}{r}, \tag{A1}$$

$$\frac{|\mathbf{p}_L|}{|\mathbf{p}_R|} = \frac{r_R}{r_L}, \tag{A2}$$

where $|\mathbf{p}_L|$ and $|\mathbf{p}_R|$ are the amplitude of sound pressures at sensors 22, 24. Fig. 12 depicts sensors 22, 24 and a representative acoustic source S1 within the range of reception to provide input signals $x_{Ln}(t)$ and $x_{Rn}(t)$. According to the geometry illustrated in Fig. 12, the distances r_L and r_R from the source S1 to the left and right sensors, respectively, are given by equations (A3) and (A4), as follows:

$$r_L = \sqrt{(l\sin\theta_i + D/2)^2 + (l\cos\theta_i)^2} = \sqrt{l^2 + lD\sin\theta_i + D^2/4}\,, \quad \text{(A3)}$$

$$r_R = \sqrt{(l\sin\theta_i - D/2)^2 + (l\cos\theta_i)^2} = \sqrt{l^2 - lD\sin\theta_i + D^2/4}\,. \quad \text{(A4)}$$

For a given delayed signal pair in the dual delay-line 342 of FIG. 11 to become equalized under this approach, the factors $\alpha_i(m)$ and $\alpha_{I-i+1}(m)$ must satisfy equation (A5) as follows:

$$|\mathbf{p}_L|\alpha_i(m) = |\mathbf{p}_R|\alpha_{I-i+1}(m)\,. \quad \text{(A5)}$$

Substituting equation (A2) into equation (A5), equation (A6) results as follows:

$$\frac{r_L}{r_R} = \frac{\alpha_i(m)}{\alpha_{I-i+1}(m)}\,. \quad \text{(A6)}$$

By defining the value of $\alpha_i(m)$ in accordance with equation (A7) as follows:

$$\alpha_i(m) = K\sqrt{l^2 + lD\sin\theta_i + D^2/4}\,, \quad \text{(A7)}$$

where, K is in units of inverse length and is chosen to provide a convenient amplitude level, the value of $\alpha_{I-i+1}(m)$ is given by equation (A8) as follows:

$$\alpha_{I-i+1}(m) = K\sqrt{l^2 + lD\sin\theta_{I-i+1} + D^2/4} = K\sqrt{l^2 - lD\sin\theta_i + D^2/4}\,, \quad \text{(A8)}$$

where, the relation $\sin\theta_{I-i+1} = -\sin\theta_i$ can be obtained by substituting I-i+1 into i in equation (3). By substituting equations (A7) and (A8) into equation (A6), it may be verified that the values assigned to $\alpha_i(m)$ in equation (A7) satisfy the condition established by equation (A6).

After obtaining the equalization factors $\alpha_i(m)$ in accordance with this embodiment, minor adjustments are preferably made to calibrate for asymmetries in the sensor arrangement and other departures from the ideal case such as those that might result from media absorption of acoustic energy, an acoustic source geometry other than a point source, and dependence of amplitude decline on parameters other than distance.

After equalization by factors $\alpha_i(m)$ with multipliers 346, the in-phase desired signal component is generally the same in the left and right channels of the dual delay lines 342 for the delayed signal pairs corresponding to $i = i_{\text{signal}} = s$, and the in-phase noise signal component is generally the same in the left and right channels of the dual

delay lines 342 for the delayed signal pairs corresponding to $i = i_{\text{noise}} = g$ for the case of a single, predominant interfering noise source. The desired signal at $i=s$ may be expressed as $S_n(m) = A_s \exp[j(\omega_m t + \phi_s)]$; and the interfering signal at $i=g$ may be expressed as $G_n(m) = A_g \exp[j(\omega_m t+\phi_g)]$, where ϕ_s and ϕ_g denote initial phases. Based on these models, equalized signals $\alpha_i(m) X_{Ln}{}^{(i)}(m)$ for the left channel and $\alpha_{I-i+1}(m) X_{Rn}{}^{(i)}(m)$ for the right channel at any arbitrary point i (except $i = s$) along dual delay lines 342 may be expressed in equations (5) and (6) as follows:

$$\alpha_i(m) X_{Ln}^{(i)}(m) = A_s \exp j[\omega_m(t+\tau_s-\tau_i)+\phi_s] + A_g \exp j[\omega_m(t+\tau_g-\tau_i)+\phi_g], \qquad (5)$$

$$\alpha_{I-i+1}(m) X_{Rn}^{(i)}(m) =$$

$$A_s \exp j[\omega_m(t+\tau_{I-s+1}-\tau_{I-i+1})+\phi_s] + A_g \exp j[\omega_m(t+\tau_{I-g+1}-\tau_{I-i+1})+\phi_g]. \qquad (6)$$

where equations (7) and (8) further define certain terms of equations (5) and (6) as follows:

$$X_{Ln}^{(i)}(m) = X_{Ln}(m) \exp(-j2\pi f_m \tau_i) \qquad (7)$$

$$X_{Rn}^{(i)}(m) = X_{Rn}(m) \exp(-j2\pi f_m \tau_{I-i+1}) \qquad (8)$$

Each signal pair $\alpha_i(m) X_{Ln}{}^{(i)}(m)$ and $\alpha_{I-i+1}(m) X_{Rn}{}^{(i)}(m)$ is input to a corresponding operation stage 354 of a corresponding one of operation arrays 352 for all m; where each operator array 352 corresponds to a different value of m as in the case of dual delay lines 342. For a given operation array 352, operation stages 354 corresponding to each value of I, except i=s, perform the operation defined by equation (9) as follows:

$$X_n^{(i)}(m) = \frac{\alpha_i(m) X_{Ln}^{(i)}(m) - \alpha_{I-i+1}(m) X_{Rn}^{(i)}(m)}{(\alpha_i/\alpha_s)\exp[j\omega_m(\tau_s-\tau_i)] - (\alpha_{I-i+1}/\alpha_{I-s+1})\exp[j\omega_m(\tau_{I-s+1}-\tau_{I-i+1})]},$$

$$\text{for } i \neq s. \qquad (9)$$

If the value of the denominator in equation (9) is too small, a small positive constant ε is added to the denominator to limit the magnitude of the output signal $X_n{}^{(i)}(m)$. No operation is performed by the operation stage 354 on the signal pair corresponding to $i=s$ for all m (all operation arrays 352 of signal operator 350).

Equation (9) is comparable to the expressions CE1 and CE2 of system 10; however, equation (9) includes equalization elements $\alpha_i(m)$ and is organized into a

single expression. With the outputs from operation array 352, the simultaneous localization and identification of the spectral content of the desired signal may be performed with system 310. Localization and extraction with system 310 are further described by the signal flow diagram of Fig. 13 and the following mathematical model. By substituting equations (5) and (6) into equation (9), equation (10) results as follows:

$$X_n^{(i)}(m) = S_n(m) + G_n(m)\cdot \upsilon_{s,g}^{(i)}(m), \qquad i \neq s \qquad (10)$$

where equation (11) further defines:

$$\upsilon_{s,g}^{(i)}(m) = \frac{(\alpha_i/\alpha_g)\exp[j\omega_m(\tau_g-\tau_i)] - (\alpha_{I-i+1}/\alpha_{I-g+1})\exp[j\omega_m(\tau_{I-g+1}-\tau_{I-i+1})]}{(\alpha_i/\alpha_s)\exp[j\omega_m(\tau_s-\tau_i)] - (\alpha_{I-i+1}/\alpha_{I-s+1})\exp[j\omega_m(\tau_{I-s+1}-\tau_{I-i+1})]}, \quad i \neq s \qquad (11)$$

By applying equation (2) to equation (11), equation (12) results as follows:

$$\upsilon_{s,g}^{(i)}(m) = \frac{(\alpha_i/\alpha_g)\exp[j\omega_m(\tau_g-\tau_i)] - (\alpha_{I-i+1}/\alpha_{I-g+1})\exp[-j\omega_m(\tau_g-\tau_i)]}{(\alpha_i/\alpha_s)\exp[j\omega_m(\tau_s-\tau_i)] - (\alpha_{I-i+1}/\alpha_{I-s+1})\exp[-j\omega_m(\tau_s-\tau_i)]}, \qquad i \neq s. \qquad (12)$$

The energy of the signal $X_n^{(i)}(m)$ is expressed in equation (13) as follows:

$$\left|X_n^{(i)}(m)\right|^2 = \left|S_n(m) + G_n(m)\cdot \upsilon_{s,g}^{(i)}(m)\right|^2. \qquad (13)$$

A signal vector may be defined:

$$\mathbf{x}^{(i)} = (X_1^{(i)}(1), X_1^{(i)}(2), \ldots, X_1^{(i)}(M), X_2^{(i)}(1), \ldots, X_2^{(i)}(M), \ldots, X_N^{(i)}(1), \ldots, X_N^{(i)}(M))^T,$$

$$i=1,\ldots,I,$$

where, T denotes transposition. The energy $\|x^{(i)}\|_2^2$ of the vector $x^{(i)}$ is given by equation (14) as follows:

$$\|\mathbf{x}^{(i)}\|_2^2 = \sum_{n=1}^{N}\sum_{m=1}^{M}\left|X_n^{(i)}(m)\right|^2 = \sum_{n=1}^{N}\sum_{m=1}^{M}\left|S_n(m) + G_n(m)\cdot \upsilon_{s,g}^{(i)}(m)\right|^2, \qquad i=1,\ldots,I. \qquad (14)$$

Equation (14) is a double summation over time and frequency that approximates a double integration in a continuous time domain representation.

Further defining the following vectors:

$\mathbf{s} = (S_1(1), S_1(2), \ldots, S_1(M), S_2(1), \ldots, S_2(M), \ldots, S_N(1), \ldots, S_N(M))^T$, and

$\mathbf{g}^{(i)} = (G_1(1)\upsilon_{s,g}^{(i)}(1), G_1(2)\upsilon_{s,g}^{(i)}(2), \ldots, G_1(M)\upsilon_{s,g}^{(i)}(M), G_2(1)\upsilon_{s,g}^{(i)}(1), \ldots, G_2(M)\upsilon_{s,g}^{(i)}(M), \ldots,$
$G_N(1)\upsilon_{s,g}^{(i)}(1), \ldots, G_N(M)\upsilon_{s,g}^{(i)}(M))^T$, where $i = 1, \ldots, I$,

the energy of vectors s and $g^{(i)}$ are respectively defined by equations (15) and (16) as follows:

$$\|\mathbf{s}\|_2^2 = \sum_{n=1}^{N}\sum_{m=1}^{M} |S_n(m)|^2 \tag{15}$$

$$\|\mathbf{g}^{(i)}\|_2^2 = \sum_{n=1}^{N}\sum_{m=1}^{M} \left|G_n(m) \cdot \upsilon_{s,g}^{(i)}(m)\right|^2, \quad i=1,\ldots,I. \tag{16}$$

For a desired signal that is independent of the interfering source, the vectors s and $g^{(i)}$ are orthogonal. In accordance with the Theorem of Pythagoras, equation (17) results as follows:

$$\|\mathbf{x}^{(i)}\|_2^2 = \|\mathbf{s} + \mathbf{g}^{(i)}\|_2^2 = \|\mathbf{s}\|_2^2 + \|\mathbf{g}^{(i)}\|_2^2, \quad i=1,\ldots,I. \tag{17}$$

Because $\|g^{(i)}\|_2^2 \geq 0$, equation (18) results as follows:

$$\|\mathbf{x}^{(i)}\|_2^2 \geq \|\mathbf{s}^{(i)}\|_2^2, \quad i=1,\ldots,I. \tag{18}$$

The equality in equation (18) is satisfied only when $\|g^{(i)}\|_2^2 = 0$, which happens if either of the following two conditions are met: (a) $G_n(m) = 0$, i.e., the noise source is silent – in which case there is no need for doing localization of the noise source and noise cancellation; and (b) $\upsilon_{s,g}^{(i)}(m) = 0$; where equation (12) indicates that this second condition arises for $i = g = i_{noise}$. Therefore, $\|x^{(i)}\|_2^2$ has its minimum at $i = g = i_{noise}$, which according to equation (18) is $\|s\|_2^2$. Equation (19) further describes this condition as follows:

$$\|\mathbf{s}\|_2^2 = \|\mathbf{x}^{(i_{noise})}\|_2^2 = \min_i \|\mathbf{x}^{(i)}\|_2^2. \tag{19}$$

Thus, the localization procedure includes finding the position i_{noise} along the operation array 352 for each of the delay lines 342 that produces the minimum value of $\|x^{(i)}\|_2^2$. Once the location i_{noise} along the dual delay line 342 is determined, the azimuth position of the noise source may be determined with equation (3). The estimated noise location i_{noise} may be utilized for noise cancellation or extraction of the desired signal as further described hereinafter. Indeed, operation stages 354 for all m corresponding to $i = i_{\text{noise}}$ provide the spectral components of the desired signal as given by equation (20):

$$\hat{S}_n(m) = X_n^{(i_{\text{noise}})}(m) = S_n(m) + G_n(m) \cdot v_{s,g}^{(i_{\text{noise}})}(m) = S_n(m)\,. \qquad (20)$$

Localization operator 360 embodies the localization technique of system 310. Fig. 13 further depicts operator 360 with coupled pairs of summation operators 362 and 364 for each value of integer index i; where i=1,...,I. Collectively, summation operators 362 and 364 perform the operation corresponding to equation (14) to generate $\|x^{(i)}\|_2^2$ for each value of i. For each transform time frame n, the summation operators 362 each receive $X_n^{(i)}(1)$ through $X_n^{(i)}(M)$ inputs from operation stages 354 corresponding to their value of i and sums over frequencies m=1 through m=M. For the illustrated example, the upper summation operator 362 corresponds to i=1 and receives signals $X_n^{(1)}(1)$ through $X_n^{(1)}(M)$ for summation; and the lower summation operator 362 corresponds to i=I and receives signals $X_n^{(I)}(1)$ through $X_n^{(I)}(M)$ for summation.

Each summation operator 364 receives the results for each transform time frame n from the summation operator 362 corresponding to the same value of i and accumulates a sum of the results over time corresponding to n=1 through n=N transform time frames; where N is a quantity of time frames empirically determined to be suitable for localization. For the illustrated example, the upper summation operator 364 corresponds to i=1 and sums the results from the upper summation operator 362 over N samples; and the lower summation operator 364 corresponds to i=I and sums the results from the lower summation operator 362 over N samples.

The I number of values of $\|x^{(i)}\|_2^2$ resulting from the I number of summation operators 364 are received by stage 366. Stage 366 compares the I number of $\|x^{(i)}\|_2^2$ values to determine the value of i corresponding to the minimum $\|x^{(i)}\|_2^2$. This value of i is output by stage 366 as $i = g = i_{\text{noise}}$.

Referring back to Fig. 10, post-localization processing by system 310 is further described. When equation (9) is applied to the pair inputs of delay lines 342 at i=g, it corresponds to the position of the off-axis noise source and equation (20) shows it provides an approximation of the desired signal $\acute{S}_n(m)$. To extract signal $\acute{S}_n(m)$, the index value i=g is sent by stage 366 of localization unit 360 to extraction operator 380. In response to g, extraction operator 380 routes the outputs $X_n^{(g)}(1)$ through $X_n^{(g)}(M) = \acute{S}_n(m)$ to Inverse Fourier Transform (IFT) stage 82 operatively coupled thereto. For this purpose, extraction operator 380 preferably includes a multiplexer or matrix switch that has IxM complex inputs and M complex outputs; where a different set of M inputs is routed to the outputs for each different value of the index I in response to the output from stage 366 of localization operator 360.

Stage 82 converts the M spectral components received from extraction unit 380 to transform the spectral approximation of the desired signal, $\acute{S}_n(m)$, from the frequency domain to the time domain as represented by signal $\acute{s}_n(k)$. Stage 82 is operatively coupled to digital-to-analog (D/A) converter 84. D/A converter 84 receives signal $\acute{s}_n(k)$ for conversion from a discrete form to an analog form represented by $\acute{s}_n(t)$. Signal $\acute{s}_n(t)$ is input to output device 90 to provide an auditory representation of the desired signal or other indicia as would occur to those skilled in the art. Stage 82, converter 84, and device 90 are further described in connection with system 10.

Another form of expression of equation (9) is given by equation (21) as follows:

$$X_n^{(i)}(m) = w_{Ln}(m) X_{Ln}^{(i)}(m) + w_{Rn}(m) X_{Rn}^{(i)}(m) . \qquad (21)$$

The terms w_{Ln} and w_{Rn} are equivalent to beamforming weights for the left and right channels, respectively. As a result, the operation of equation (9) may be equivalently modeled as a beamforming procedure that places a null at the location corresponding to the predominant noise source, while steering to the desired output signal $\acute{s}_n(t)$.

Fig. 14 depicts system 410 of still another embodiment of the present invention. System 410 is depicted with several reference numerals that are the same as those used in connection with systems 10 and 310 and are intended to designate like features. A number of acoustic sources 412, 414, 416, 418 are depicted in Fig. 14 within the reception range of acoustic sensors 22, 24 of system 410. The positions of

sources 412, 414, 416, 418 are also represented by the azimuth angles relative to axis AZ that are designated with reference numerals 412a, 414a, 416a, 418a. As depicted, angles 412a, 414a, 416a, 418a correspond to about 0°, +20°, +75°, and -75°, respectively. Sensors 22, 24 are operatively coupled to signal processor 430 with axis AZ extending about midway therebetween. Processor 430 receives input signals $x_{Ln}(t)$, $x_{Rn}(t)$ from sensors 22, 24 corresponding to left channel L and right channel R as described in connection with system 310. Processor 430 processes signals $x_{Ln}(t)$, $x_{Rn}(t)$ and provides corresponding output signals to output devices 90, 490 operatively coupled thereto.

Referring additionally to the signal flow diagram of Fig. 15, selected features of system 410 are further illustrated. System 410 includes D/A converters 34a, 34b and DFT stages 36a, 36b to provide the same left and right channel processing as described in connection with system 310. System 410 includes delay operator 340 and signal operator 350 as described for system 310; however it is preferred that equalization factors $\alpha_i(m)$ (i=1, ..., I) be set to unity for the localization processes associated with localization operator 460 of system 410. Furthermore, localization operator 460 of system 410 directly receives the output signals of delay operator 340 instead of the output signals of signal operator 350, unlike system 310.

The localization technique embodied in operator 460 begins by establishing two-dimensional (2-D) plots of coincidence loci in terms of frequency versus azimuth position. The coincidence points of each loci represent a minimum difference between the left and right channels for each frequency as indexed by m. This minimum difference may be expressed as the minimum magnitude difference $\delta X_n^{(i)}(m)$ between the frequency domain representations $X_{Lp}^{(i)}(m)$ and $X_{Lp}^{(i)}(m)$, at each discrete frequency m, yielding M/2 potentially different loci. If the acoustic sources are spatially coherent, then these loci will be the same across all frequencies. This operation is described in equations (22)-(25) as follows:

$$i_n(m) = \arg\min_i \{\delta X_n^{(i)}(m)\}, \qquad m=1, \ldots, M/2. \qquad (22)$$

$$\delta X_n^{(i)}(m) = \left| X_{Ln}^{(i)}(m) - X_{Rn}^{(i)}(m) \right|, \qquad i=1, \ldots, I;\ m=1, \ldots, M/2, \qquad (23)$$

$$X_{Ln}^{(i)}(m) = X_{Ln}(m)\exp(-j2\pi\tau_i m / M), \qquad i=1, \ldots, I;\ m=1, \ldots, M/2, \qquad (24)$$

$$X_{Rn}^{(i)}(m) = X_{Rn}(m)\exp(-j2\pi\tau_{I-i+1} m / M), \qquad i=1, \ldots, I;\ m=1, \ldots, M/2. \qquad (25)$$

If the amplitudes of the left and right channels are generally the same at a given position along dual delay lines 342 of system 410 as indexed by i, then the values of $\delta X_n^{(i)}(m)$ for the corresponding value of i is minimized, if not essentially zero. It is noted that, despite inter-sensor intensity differences, equalization factors $\alpha_i(m)$ (i=1, ..., I) should be maintained close to unity for the purpose of coincidence detection; otherwise, the minimal $\delta X_n^{(i)}(m)$ will not correspond to the in-phase (coincidence) locations.

An alternative approach may be based on identifying coincidence loci from the phase difference. For this phase difference approach, the minimum of the phase difference between the left and right channel signals at positions along the dual delay lines 342, as indexed by i, are located as described by the following equations (26) and (27):

$$i_n(m) = \arg\min_i \{\delta X_n^{(i)}(m)\}, \qquad m=1, \ldots, M/2, \qquad (26)$$

$$\delta X_n^{(i)}(m) = \left|\mathrm{Im}[X_{Ln}^{(i)}(m) X_{Rn}^{(i)}(m)^{\dagger}]\right|, \quad i=1, \ldots, I;\ m=1, \ldots, M/2, \qquad (27)$$

where, Im[•] denotes the imaginary part of the argument, and the superscript † denotes a complex conjugate. Since the phase difference technique detects the minimum angle between two complex vectors, there is also no need to compensate for the inter-sensor intensity difference.

While either the magnitude or phase difference approach may be effective without further processing to localize a single source, multiple sources often emit spectrally overlapping signals that lead to coincidence loci which correspond to nonexistent or phantom sources (e.g., at the midpoint between two equal intensity sources at the same frequency). Fig. 17 illustrates a 2-D coincidence plot 500 in terms of frequency in Hertz (Hz) along the vertical axis and azimuth position in degrees along the horizontal axis. Plot 500 indicates two sources corresponding to the generally vertically aligned locus 512a at about –20 degrees and the vertically aligned locus 512b at about + 40 degrees. Plot 500 also includes misidentified or phantom source points 514a, 514b, 514c, 514d, 514e at other azimuths positions that correspond to frequencies where both sources have significant energy. Plots having more than two differently located competing acoustic sources generally result in an even more complex plot.

To reduce the occurrence of phantom information in the 2-D coincidence plot data, localization operator 460 integrates over time and frequency. When the signals are not correlated at each frequency, the mutual interference between the signals can be gradually attenuated by the temporal integration. This approach averages the locations of the coincidences, not the value of the function used to determine the minima, which is equivalent to applying a Kronecker delta function, $\delta(i-i_n(m))$ to $\delta X_n^{(i)}(m)$ and averaging the $\delta(i-i_n(m))$ over time. In turn, the coincidence loci corresponding to the true position of the sources are enhanced. Integration over time applies a forgetting average to the 2-D coincidence plots acquired over a predetermined set of transform time frames from $n = 1, \ldots, N$; and is expressed by the summation approximation of equation (28) as follows:

$$P_N(\theta_i, m) = \sum_{n=1}^{N} \beta^{N-n} \delta(i - i_n(m)), \qquad i=1, \ldots, I;\ m=1, \ldots, M/2, \tag{28}$$

where, $0 < \beta < 1$ is a weighting coefficient which exponentially de-emphasizes (or forgets) the effect of previous coincidence results, $\delta(\bullet)$ is the Kronecker delta function, θ_i represents the position along the dual delay-lines 342 corresponding to spatial azimuth θ_i [equation (2)], and N refers to the current time frame. To reduce the cluttering effect due to instantaneous interactions of the acoustic sources, the results of equation (28) are tested in accordance with the relationship defined by equation (29) as follows:

$$P_N(\theta_i, m) = \begin{cases} P_N(\theta_i, m), & P_N(\theta_i, m) \geq \Gamma \\ 0, & \text{otherwise.} \end{cases} \tag{29}$$

where $\Gamma \geq 0$, is an empirically determined threshold. While this approach assumes the inter-sensor delays are independent of frequency, it has been found that departures from this assumption may generally be considered negligible.

By integrating the coincidence plots across frequency, a more robust and reliable indication of the locations of sources in space is obtained. Integration of $P_n(\theta_i, m)$ over frequency produces a localization pattern which is a function of azimuth. Two techniques to estimate the true position of the acoustic sources may be utilized. The first estimation technique is solely based on the straight vertical traces

across frequency that correspond to different azimuths. For this technique, θ_d denotes the azimuth with which the integration is associated, such that $\theta_d = \theta_i$, and results in the summation over frequency of equation (30) as follows:

$$H_N(\theta_d) = \sum_m P_N(\theta_d, m), \qquad d=1, \ldots, I. \qquad (30)$$

where, equation (30) approximates integration over time.

The peaks in $H_n(\theta_d)$ represent the source azimuth positions. If there are Q sources, Q peaks in $H_N(\theta_d)$ may generally be expected. When compared with the patterns $\delta(i-i_n(m))$ at each frequency, not only is the accuracy of localization enhanced when more than one sound source is present, but also almost immediate localization of multiple sources for the current frame is possible. Furthermore, although a dominant source usually has a higher peak in $H_N(\theta_d)$ than do weaker sources, the height of a peak in $H_N(\theta_d)$ only indirectly reflects the energy of the sound source. Rather, the height is influenced by several factors such as the energy of the signal component corresponding to θ_d relative to the energy of the other signal components for each frequency band, the number of frequency bands, and the duration over which the signal is dominant. In fact, each frequency is weighted equally in equation (28). As a result, masking of weaker sources by a dominant source is reduced. In contrast, existing time-domain cross-correlation methods incorporate the signal intensity, more heavily biasing sensitivity to the dominant source.

Notably, the interaural time difference is ambiguous for high frequency sounds where the acoustic wavelengths are less than the separation distance D between sensors 22, 24. This ambiguity arises from the occurrence of phase multiples above this inter-sensor distance related frequency, such that a particular phase difference $\Delta\phi$ cannot be distinguished from $\Delta\phi + 2\pi$. As a result, there is not a one-to-one relationship of position versus frequency above a certain frequency. Thus, in addition to the primary vertical trace corresponding to $\theta_d = \theta_i$, there are also secondary relationships that characterize the variation of position with frequency for each ambiguous phase multiple. These secondary relationships are taken into account for the second estimation technique for integrating over frequency. Equation (31) provides a means to determine a predictive coincidence pattern for a given azimuth that accounts for these secondary relationships as follows:

$$\sin\theta_i - \sin\theta_d = \frac{\gamma_{m,d}}{\mathrm{ITD}_{max} f_m}, \qquad (31)$$

where the parameter $\gamma_{m,d}$ is an integer, and each value of $\gamma_{m,d}$ defines a contour in the pattern $P_N(\theta_i, m)$. The primary relationship is associated with $\gamma_{m,d} = 0$. For a specific θ_d, the range of valid $\gamma_{m,d}$ is given by equation (32) as follows:

$$-\mathrm{ITD}_{max} f_m (1 + \sin\theta_d) \le \gamma_{m,d} \le \mathrm{ITD}_{max} f_m (1 - \sin\theta_d) \qquad (32)$$

The graph 600 of Fig. 18 illustrates a number of representative coincidence patterns 612, 614, 616, 618 determined in accordance with equations (31) and (32); where the vertical axis represents frequency in Hz and the horizontal axis represents azimuth position in degrees. Pattern 612 corresponds to the azimuth position of 0°. Pattern 612 has a primary relationship corresponding to the generally straight, solid vertical line 612a and a number of secondary relationships corresponding to curved solid line segments 612b. Similarly, patterns 614, 616, 618 correspond to azimuth positions of -75°, 20°, and 75° and have primary relationships shown as straight vertical lines 614a, 616a, 618a and secondary relationships shown as curved line segments 614b, 616b, 618b, in correspondingly different broken line formats. In general, the vertical lines are designated primary contours and the curved line segments are designated secondary contours. Coincidence patterns for other azimuth positions may be determined with equations (31) and (32) as would occur to those skilled in the art.

Notably, the existence of these ambiguities in $P_N(\theta_i, m)$ may generate artifactual peaks in $H_N(\theta_d)$ after integration along $\theta_d = \theta_i$. Superposition of the curved traces corresponding to several sources may induce a noisier $H_N(\theta_d)$ term. When far away from the peaks of any real sources, the artifact peaks may erroneously indicate the detection of nonexistent sources; however, when close to the peaks corresponding to true sources, they may affect both the detection and localization of peaks of real sources in $H_N(\theta_d)$. When it is desired to reduce the adverse impact of phase ambiguity, localization may take into account the secondary relationships in addition to the primary relationship for each given azimuth position. Thus, a coincidence pattern for each azimuthal direction θ_d ($d=1, \ldots, I$) of interest may be determined and

plotted that may be utilized as a "stencil" window having a shape defined by $P_N(\theta_i, m)$ (i=1, ..., I; m=1, ..., M). In other words, each stencil is a predictive pattern of the coincidence points attributable to an acoustic source at the azimuth position of the primary contour, including phantom loci corresponding to other azimuth positions as a factor of frequency. The stencil pattern may be used to filter the data at different values of m.

By employing the equation (32), the integration approximation of equation (30) is modified as reflected in the following equation (33):

$$H_N(\theta_d) = \frac{1}{A(\theta_d)} \sum_m P_N[\sin^{-1}(\frac{\gamma_{m,d}}{\mathrm{ITD}_{\max} f_m} + \sin\theta_d), m], \qquad d=1, \ldots, I, \qquad (33)$$

where $A(\theta_d)$ denotes the number of points involved in the summation. Notably, equation (30) is a special case of equation (33) corresponding to $\gamma_{m,d}=0$. Thus, equation (33) is used in place of equation (30) when the second technique of integration over frequency is desired.

As shown in equation (2), both variables θ_i and τ_i are equivalent and represent the position in the dual delay-line. The difference between these variables is that θ_i indicates location along the dual delay-line by using its corresponding spatial azimuth, whereas τ_i denotes location by using the corresponding time-delay unit of value τ_i. Therefore, the stencil pattern becomes much simpler if the stencil filter function is expressed with τ_i as defined in the following equation (34):

$$\tau_i - \tau_d = \frac{\gamma_{m,d}}{2f_m}, \qquad (34)$$

where, τ_d relates to θ_d through equation (4). For a specific τ_d, the range of valid $\gamma_{m,d}$ is given by equation (35) as follows:

$$-(\mathrm{ITD}_{\max}/2 + \tau_d) f_m \le \gamma_{m,d} \le (\mathrm{ITD}_{\max}/2 - \tau_d) f_m, \ \gamma_{m,d} \text{ is an integer.} \qquad (35)$$

Changing value of τ_d only shifts the coincidence pattern (or stencil pattern) along the τ_i-axis without changing its shape. The approach characterized by equations (34) and

(35) may be utilized as an alternative to separate patterns for each azimuth position of interest; however, because the scaling of the delay units τ_i is uniform along the dual delay-line, azimuthal partitioning by the dual delay-line is not uniform, with the regions close to the median plane having higher azimuthal resolution. On the other hand, in order to obtain an equivalent resolution in azimuth, using a uniform τ_i would require a much larger I of delay units than using a uniform θ_i.

The signal flow diagram of Fig. 16 further illustrates selected details concerning localization operator 460. With equalization factors $\alpha_i(m)$ set to unity, the delayed signal of pairs of delay stages 344 are sent to coincidence detection operators 462 for each frequency indexed to m to determine the coincidence points. Detection operators 462 determine the minima in accordance with equation (22) or (26). Each coincidence detection operator 462 sends the results $i_n(m)$ to a corresponding pattern generator 464 for the given m. Generators 464 build a 2-D coincidence plot for each frequency indexed to m and pass the results to a corresponding summation operator 466 to perform the operation expressed in equation (28) for that given frequency. Summation operators 466 approximate integration over time. In Fig. 16, only operators 462, 464, and 466 corresponding to m =1 and m =M are illustrated to preserve clarity, with those corresponding to m =2 through m = M-1 being represented by ellipses.

Summation operators 466 pass results to summation operator 468 to approximate integration over frequency. Operators 468 may be configured in accordance with equation (30) if artifacts resulting from the secondary relationships at high frequencies are not present or may be ignored. Alternatively, stencil filtering with predictive coincidence patterns that include the secondary relationships may be performed by applying equation (33) with summation operator 468.

Referring back to Fig. 15, operator 468 outputs $H_N(\theta_d)$ to output device 490 to map corresponding acoustic source positional information. Device 490 preferably includes a display or printer capable of providing a map representative of the spatial arrangement of the acoustic sources relative to the predetermined azimuth positions. In addition, the acoustic sources may be localized and tracked dynamically as they move in space. Movement trajectories may be estimated from the sets of locations $\delta(i-i_n(m))$ computed at each sample window n. For other embodiments incorporating system 410 into a small portable unit, such as a hearing aid, output device 490 is

preferably not included. In still other embodiments, output device 90 may not be included.

The localization techniques of localization operator 460 are particularly suited to localize more than two acoustic sources of comparable sound pressure levels and frequency ranges, and need not specify an on-axis desired source. As such, the localization techniques of system 410 provide independent capabilities to localize and map more than two acoustic sources relative to a number of positions as defined with respect to sensors 22, 24. However, in other embodiments, the localization capability of localization operator 460 may also be utilized in conjunction with a designated reference source to perform extraction and noise suppression. Indeed, extraction operator 480 of the illustrated embodiment incorporates such features as more fully described hereinafter.

Existing systems based on a two sensor detection arrangement generally only attempt to suppress noise attributed to the most dominant interfering source through beamforming. Unfortunately, this approach is of limited value when there are a number of comparable interfering sources at proximal locations.

It has been discovered that by suppressing one or more different frequency components in each of a plurality of interfering sources after localization, it is possible to reduce the interference from the noise sources in complex acoustic environments, such as in the case of multi-talkers, in spite of the temporal and frequency overlaps between talkers. Although a given frequency component or set of components may only be suppressed in one of the interfering sources for a given time frame, the dynamic allocation of suppression of each of the frequencies among the localized interfering acoustic sources generally results in better intelligibility of the desired signal than is possible by simply nulling only the most offensive source at all frequencies.

Extraction operator 480 provides one implementation of this approach by utilizing localization information from localization operator 460 to identify Q interfering noise sources corresponding to positions other than $i = s$. The positions of the Q noise sources are represented by i=*noise1*, *noise2*,..., *noiseQ*. Notably, operator 480 receives the outputs of signal operator 350 as described in connection with system 310, that presents corresponding signals $X_n^{(i=\text{noise1})}(m)$, $X_n^{(i=noise2)}(m)$, ..., $X_n^{(i=noiseQ)}(m)$ for each frequency m. These signals include a component of the desired signal at frequency m as well as components from sources other than the one to be

canceled. For the purpose of extraction and suppression, the equalization factors $\alpha_i(m)$ need not be set to unity once localization has taken place. To determine which frequency component or set of components to suppress in a particular noise source, the amplitudes of $X_n^{(i=\text{noise1})}(m)$, $X_n^{(i=\text{noise2})}(m)$, ..., $X_n^{(i=noiseQ)}(m)$ are calculated and compared. The minimum $X_n^{(inoise)}(m)$, is taken as output $\acute{S}_n(m)$ as defined by the following equation (36):

$$\acute{S}_n(m) = X_n^{(inoise)}(m), \tag{36}$$

where, $X_n^{(inoise)}(m)$ satisfies the condition expressed by equation (37) as follows:

$$|X_n^{(inoise)}(m)| = \min\{|X_n^{(i=\text{noise1})}(m)|, |X_n^{(i=\text{noise2})}(m)|, \ldots, |X_n^{(i=noiseQ)}(m)|,$$

$$|\alpha_s(m)\, X_{Ln}^{(s)}(m)|\}; \tag{37}$$

for each value of m. It should be noted that, in equation (37), the original signal $\alpha_s(m)\, X_{Ln}^{(s)}(m)$ is included. The resulting beam pattern may at times amplify other less intense noise sources. When the amount of noise amplification is larger than the amount of cancellation of the most intense noise source, further conditions may be included in operator 480 to prevent changing the input signal for that frequency at that moment.

Processors 30, 330, 430 include one or more components that embody the corresponding algorithms, stages, operators, converters, generators, arrays, procedures, processes, and techniques described in the respective equations and signal flow diagrams in software, hardware, or both utilizing techniques known to those skilled in the art. Processors 30, 330, 430 may be of any type as would occur to those skilled in the art; however, it is preferred that processors 30, 330, 430 each be based on a solid-state, integrated digital signal processor with dedicated hardware to perform the necessary operations with a minimum of other components.

Systems 310, 410 may be sized and adapted for application as a hearing aide of the type described in connection with Fig. 4A. In a further hearing aid embodiment, sensors application 22, 24 are sized and shaped to fit in the pinnae of a listener, and the processor algorithms are adjusted to account for shadowing caused by the head and torso. This adjustment may be provided by deriving a Head-Related-Transfer-Function (HRTF) specific to the listener or from a population average using techniques known to those skilled in the art. This function is then used to provide

appropriate weightings of the dual delay stage output signals that compensate for shadowing.

In yet another embodiment, system 310, 410 are adapted to voice recognition systems of the type described in connection with Fig. 4B. In still other embodiments, systems 310, 410 may be utilized in sound source mapping applications, or as would otherwise occur to those skilled in the art.

It is contemplated that various signal flow operators, converters, functional blocks, generators, units, stages, processes, and techniques may be altered, rearranged, substituted, deleted, duplicated, combined or added as would occur to those skilled in the art without departing from the spirit of the present inventions.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference, including, but not limited to U.S. Patent Application Serial No. 08/666,757 filed on 19 June 1996.

EXPERIMENTAL SECTION

The following experimental results are provided as nonlimited examples, and should not be construed to restrict the scope of the present invention.

EXAMPLE ONE

A Sun Sparc-20 workstation was programmed to emulate the signal extraction process of the present invention. One loudspeaker (L1) was used to emit a speech signal and another loudspeaker (L2) was used to emit babble noise in a semi-anechoic room. Two microphones of a conventional type were positioned in the room and operatively coupled to the workstation. The microphones had an inter-microphone distance of about 15 centimeters and were positioned about 3 feet from L1. L1 was aligned with the midpoint between the microphones to define a zero degree azimuth. L2 was placed at different azimuths relative to L1 approximately equidistant to the midpoint between L1 and L2.

Referring to FIG. 5, a clean speech of a sentence about two seconds long is depicted, emanating from L1 without interference from L2. FIG. 6 depicts a composite signal from L1 and L2. The composite signal includes babble noise from L2 combined with the speech signal depicted in FIG. 5. The babble noise and speech signal are of generally equal intensity (0dB) with L2 placed at a 60 degree azimuth relative to L1. FIG. 7 depicts the signal recovered from the composite signal of FIG. 6. This signal is nearly the same as the signal of FIG. 5.

FIG. 8 depicts another composite signal where the babble noise is 30dB more intense than the desired signal of FIG. 5. Furthermore, L2 is placed at only a 2 degree azimuth relative to L1. FIG. 9 depicts the signal recovered from the composite signal of FIG. 8, providing a clearly intelligible representation of the signal of FIG. 5 despite the greater intensity of the babble noise from L2 and the nearby location.

EXAMPLE TWO

Experiments corresponding to system 410 were conducted with two groups having four talkers (2 male, 2 female) in each group. Five different tests were conducted for each group with different spatial configurations of the sources in each test. The four talkers were arranged in correspondence with sources 412, 414, 416, 418 of Fig. 14 with different values for angles 412a, 414a, 416a, and 418a in each test. The illustration in Fig. 14 most closely corresponds to the first test with angle 418a being –75 degrees , angle 412a being 0 degrees, angle 414a being +20 degrees, and angle 416a being +75 degrees. The coincident patterns 612, 614, 616, and 618 of Fig. 18 also correspond to the azimuth positions of –75 degrees, 0 degrees, +20 degrees, and +75 degrees.

The experimental set-up for the tests utilized two microphones for sensors 22, 24 with an inter-microphone distance of about 144mm. No diffraction or shadowing effect existed between the two microphones, and the inter-microphone intensity difference was set to zero for the tests. The signals were low-pass filtered at 6 kHz and sampled at a 12.8-kHz rate with 16-bit quantization. A Wintel-based computer was programmed to receive the quantized signals for processing in accordance with the present invention and output the test results described hereinafter. In the short-term spectral analysis, a 20-ms segment of signal was weighted by a Hamming window and then padded with zeros to 2048 points for DFT, and thus the frequency resolution was about 6Hz. The values of the time delay units τ_i (i=1, ..., I) were determined such that the azimuth resolution of the dual delay-line was 0.5° uniformly, namely I=361. The dual delay-line used in the tests was azimuth-uniform. The coincidence detection method was based on minimum magnitude differences.

Each of the five tests consisted of four subtests in which a different talker was taken as the desired source. To test the system performance under the most difficult experimental constraint, the speech materials (four equally-intense spondaic words) were intentionally aligned temporally. The speech material was presented in free-field. The localization of the talkers was done using both the equation (30) and equation (33) techniques.

The system performance was evaluated using an objective intelligibility-weighted measure, as proposed in Peterson, P.M., "Adaptive array processing for multiple microphone hearing aids," Ph.D. Dissertation, Dept. Elect. Eng. and Comp. Sci., MIT; Res. Lab. Elect. Tech. Rept. 541, MIT, Cambridge, MA (1989). and described in detail in Liu, C. and Sideman, S., "Simulation of fixed microphone arrays for directional hearing aids," J. Acoust. Soc. Am. 100, 848-856 (1996). Specifically, intelligibility-weighted signal

cancellation, intelligibility-weighted noise cancellation, and net intelligibility-weighted gain were used.

The experimental results are presented in Tables I, II, III, and IV of FIGs. 19-22, respectively. The five tests described in Table I of FIG. 19 approximate integration over frequency by utilizing equation (30); and includes two male speakers M1, M2 and two female speakers F1, F2. The five tests described in Table II of FIG. 20 are the same as Table I, except that integration over frequency was approximated by equation (33). The five tests described in Table III of FIG. 21 approximate integration over frequency by utilizing equation (30); and includes two different male speakers M3, M4 and two different female speakers F3, F4. The five tests described in Table IV of FIG. 22 are the same as Table III, except that integration over frequency was approximated by equation (33).

For each test, the data was arranged in a matrix with the numbers on the diagonal line representing the degree of noise cancellation in dB of the desired source (ideally 0 dB) and the numbers elsewhere representing the degree of noise cancellation for each noise source. The next to the last column shows a degree of cancellation of all the noise sources lumped together, while the last column gives the net intelligibility-weighted improvement (which considers both noise cancellation and loss in the desired signal).

The results generally show cancellation in the intelligibility-weighted measure in a range of about 3~11 dB, while degradation of the desired source was generally less than about 0.1 dB). The total noise cancellation was in the range of about 8~12 dB. Comparison of the various Tables suggests very little dependence on the talker or the speech materials used in the tests. Similar results were obtained from six-talker experiments. Generally, a 7~10 dB enhancement in the intelligibility-weighted signal-to-noise ratio resulted when there were six equally loud, temporally aligned speech sounds originating from six different loudspeakers.

While the invention has been illustrated and described in detail in the drawings and foregoing description, the same is to be considered as illustrative and not restrictive in character, it being understood that only the preferred embodiment has been shown and described and that all changes, modifications, and equivalents that come within the spirit of the invention defined by the following claims are desired to be protected.

CLAIMS

What is claimed is:

1. A method, comprising:

providing a first signal from a first acoustic sensor and a second signal from a second acoustic sensor spaced apart from the first acoustic sensor, the first signal and the second signal each corresponding to two or more acoustic sources, said acoustic sources including a plurality of interfering sources and a desired source;

localizing the interfering sources from the first and second signals to provide a corresponding number of interfering source signals each corresponding to a different one of the interfering sources and each including a pluarlity of frequency components, the components each corresponding to a different frequency; and

suppressing one or more different frequency components of each of the interfering source signals to reduce noise.

2. The method of claim 1, wherein said suppressing includes extracting a desired signal representative of the desired source.

3. The method of claim 2, wherein said extracting includes determining a minimum value as a function of the interfering signals.

4. The method of claim 1, wherein said localizing includes filtering with a number of coincidence patterns each corresponding to one of a number of predetermined spatial positions relative to the first and second sensors, the patterns each providing phantom position information that varies with frequency relative to the one of the predetermined spatial positions.

5. The method of claim 1, further comprising delaying the first and second signals with a different dual delay line for each of a number of frequencies to provide a corresponding number of delayed signals to perform said localizing.

6. The method of claim 5, further comprising processing the delayed signals after said localizing to perform said suppressing.

7. The method of claim 6, further comprising:

transforming the first and second signals from a time domain form to a frequency domain form in terms of the frequencies before said delaying;

extracting a desired signal representative of the desired source, said extracting including said suppressing;

transforming the desired signal from a frequency domain form to a time domain form; and

generating an acoustic output representative of the desired source from the time domain form of the desired signal.

8. The method of claim 5, wherein the interfering signals are each determined from a unique pair of the delayed signals as a ratio between a difference in magnitude of the unique pair of the delayed signals and a difference determined as a function of an amount of delay associated with each member of the unique pair of the delayed signals.

9. A system, comprising:

a pair of spaced apart acoustic sensors each arranged to detect two or more differently located acoustic sources and correspondingly generate a pair of input signals, said acoustic sources including a desired source and a plurality of interfering sources;

a delay operator responsive to said input signals to generate a number of delayed signals therefrom;

a localization operator responsive to said delayed signals to localize said interfering sources relative to location of said sensors and provide a plurality of interfering source signals each representative of a corresponding one of said interfering sources, said interfering source signals each being represented in terms of a plurality of frequency components, said components each corresponding to a different frequency;

an extraction operator responsive to said interfering source signals to suppress at least one of said frequency components of each of said interfering source signals and extract a desired signal corresponding to said desired source, said at least one of said frequency components being different for each of said interfering source signals; and

an output device responsive to said desired signal to provide an output corresponding to said desired source.

10. The system of claim 9, wherein said localization operator includes a filter to localize said interfering sources relative to a number of positions, said filter being based on a different coincidence pattern of ambiguous positional information that varies with frequency for each of said positions.

11. The system of claim 9, further comprising:

an analog-to-digital converter responsive to said input signals to convert each of said input signals from an analog form to a digital form;

a first transformation stage responsive to said digital form of said input signals to transform said input signals from a time domain form to a frequency domain form in terms of a plurality of discrete frequencies, said delay operator including a dual delay line for each of the frequencies;

a second transformation stage responsive to said desired signal to transform said desired signal from a digital frequency domain form to a digital time domain form; and

a digital-to-analog converter responsive to said digital time domain form to convert said desired signal to an analog output form for said output device.

12. The system of claim 9, wherein said delay operator, said localization operator, and said extraction operator are provided by a solid state signal processing device.

13. The system of claim 9, wherein said desired source signal is determined as a function of said interfering signals.

14. The system of claim 9, wherein said interfering source signals are each determined from a unique pair of said delayed signals.

15. The system of claim 14, wherein said interfering signals each correspond to a ratio between a difference in magnitude of said unique pair of said delayed signals and a difference determined as a function of an amount of delay associated with each member of said unique pair of said delayed signals.

16. The system of claim 9, wherein said output device is configured to provide an acoustic output representative of said desired source.

17. A method, comprising:

positioning a first acoustic sensor and a second acoustic sensor to detect a plurality of differently located acoustic sources;

generating a first signal corresponding to said sources with said first sensor and a second signal corresponding to said sources with said second sensor;

providing a number of delayed signal pairs from the first and second signals, the delayed signal pairs each corresponding to one of a number of positions relative to the first and second sensors; and

localizing the sources as a function of the delayed signal pairs and a number of coincidence patterns, the patterns each corresponding to one of the positions and establishing an expected variation of acoustic source position information with frequency attributable to a source at the one of the positions.

18. The method of claim 17, wherein the coincidence patterns each correspond to a number of relationships characterizing a variation of phantom acoustic source position with frequency, the relationships each corresponding to a different ambiguous phase multiple.

19. The method of claim 18, further comprising determining the relationships for each of the coincidence patterns as a function of distance separating the first and second sensors.

20. The method of claim 18, wherein the relationships each correspond to a secondary contour that curves in relation to a primary contour, the primary contour representing frequency invariant acoustic source position information determined from the delayed signal pair corresponding to the one of the positions.

21. The method of claim 17, wherein said localizing includes filtering with the coincidence patterns to enhance true position information with phantom position information.

22. The method of claim 21, wherein said localizing includes integrating over time and integrating over frequency.

23. The method of claim 17, wherein the first sensor and second sensor are part of a hearing aid device and further comprising adjusting the delayed signal pairs with a head-related-transfer function.

24. The method of claim 17, further comprising:

extracting a desired signal after said localizing; and

suppressing a different set of frequency components for each of a selected number of the sources to reduce noise.

25. The method of claim 17, wherein the positions each correspond to an azimuth established relative to the first and second sensors and further comprising generating a map showing relative location of each of the sources.

26. A system, comprising:

a pair of spaced apart acoustic sensors each configured to generate a corresponding one of a pair of inputs signals, the signals being representative of a number of differently located acoustic sources;

a delay operator responsive to said input signals to generate a number of delayed signals each corresponding to one of a number of positions relative to said sensors;

a localization operator responsive to said delayed signals to determine a number of sound source localization signals from said delayed signals and a number of coincidence patterns, said patterns each corresponding to one of said positions and relating frequency varying sound source position information caused by ambiguous phase multiples to said one of said positions to improve sound source localization; and

an output device responsive to said localization signals to provide an output corresponding to at least one of said sources.

27. The system of claim 26, further comprising:

an analog-to-digital converter responsive to said input signals to convert each of said input signals from an analog form to a digital form; and

a first transformation stage responsive to said digital form of said input signals to transform said input signals from a time domain form to a frequency domain form in terms

of a plurality of discrete frequencies, said delay operator including a dual delay line for each of the frequencies.

28. The system of claim 27, further comprising:

an extraction operator responsive to said localization signals to extract a desired signal;

a second transformation stage responsive to said desired signal to transform said desired signal from a digital frequency domain form to a digital time domain form; and

a digital to analog converter responsive to said digital time domain form to convert said desired signal to an analog output form for said output device.

29. The system of claim 26, wherein said output device is configured to provide a map of acoustic source locations.

30. The system of claim 26, wherein said delay operator and said localization operator are defined by an integrated solid state signal processor.

31. The system of claim 26, wherein said localization operator responds to said delay signals to determine a closest one of said positions for one of said sources as a function of at least one of said delayed signals corresponding to said closest one of said positions and at least two other of said delayed signals corresponding to other of said positions, said at least two other of said delayed signals being determined with a corresponding one of said coincidence patterns.

32. A system, comprising:

a pair of spaced apart acoustic sensors each generating a corresponding one of a pair of inputs signals, the signals each being representative of a number of differently located sound sources;

a signal processor responsive to said sensors, said processor including: (a) a means for providing a number of delayed signals from said input signals, the delayed signals each corresponding to one of a number of positions relative to said first and second sensors; (b) a means for localizing each of said sound sources to one of said positions as a function of said delayed signals and a corresponding one of a number of patterns of frequency invariant data corresponding to one of said positions and frequency dependent data corresponding to at

least two other of said positions; (c) a means for suppressing a different frequency component of each of a selected number of said sources causing interference and for extracting a desired signal representative of one of said sources; and

an output device responsive to said desired signal to provide an output corresponding to said one of said sources.

33. The system of claim 32, wherein said processor includes a means for adjusting said delayed signals with a head-related-transfer-function.

ABSTRACT OF THE DISCLOSURE

A desired acoustic signal is extracted from a noisy environment by generating a signal representative of the desired signal with a processor. The processor receives aural signals from two sensors each at a different location. The two inputs to the processor are converted from analog to digital format and then submitted to a discrete Fourier transform process to generate discrete spectral signal representations. The spectral signals are delayed by a number of time intervals in a dual delay line to provide a number of intermediate signals, each corresponding to a different spatial location relative to the two sensors. Locations of the noise source and the desired source are determined and the spectral content of the desired signal is determined from the intermediate signal corresponding to the noise source locations. Inverse transformation of the selected intermediate signal followed by digital to analog conversion provides an output signal representative of the desired signal. Techniques to localize multiple acoustic sources are also disclosed. Further, a technique to enhance noise reduction from multiple sources based on two-sensor reception is described.

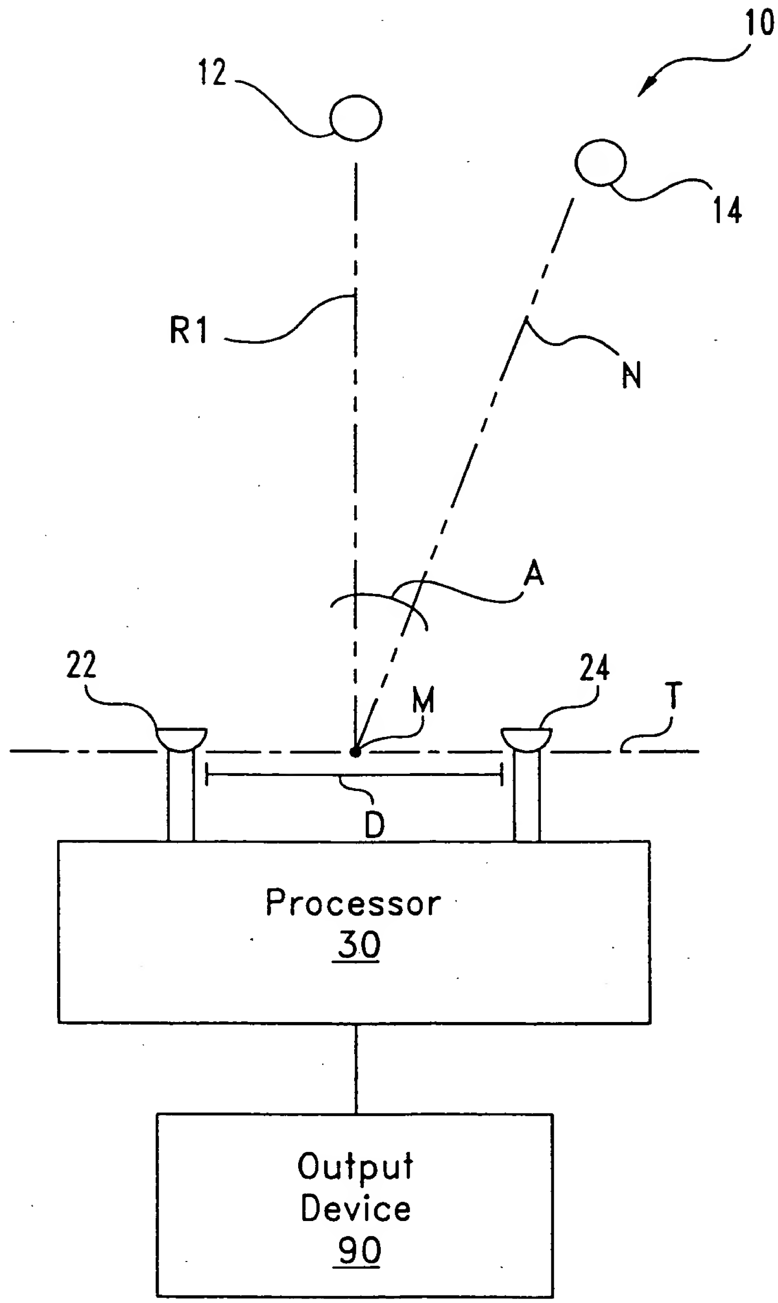


Fig. 1

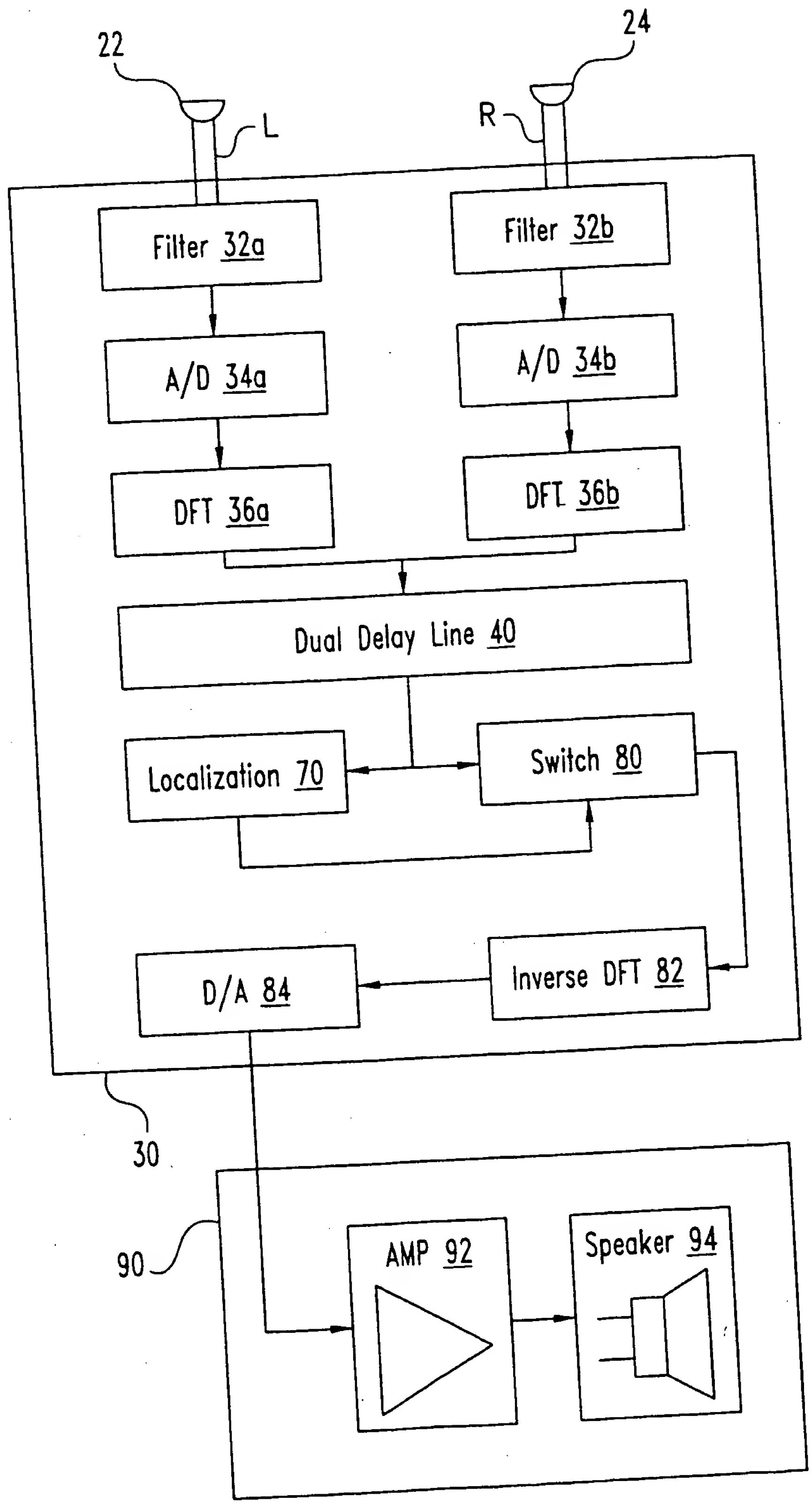
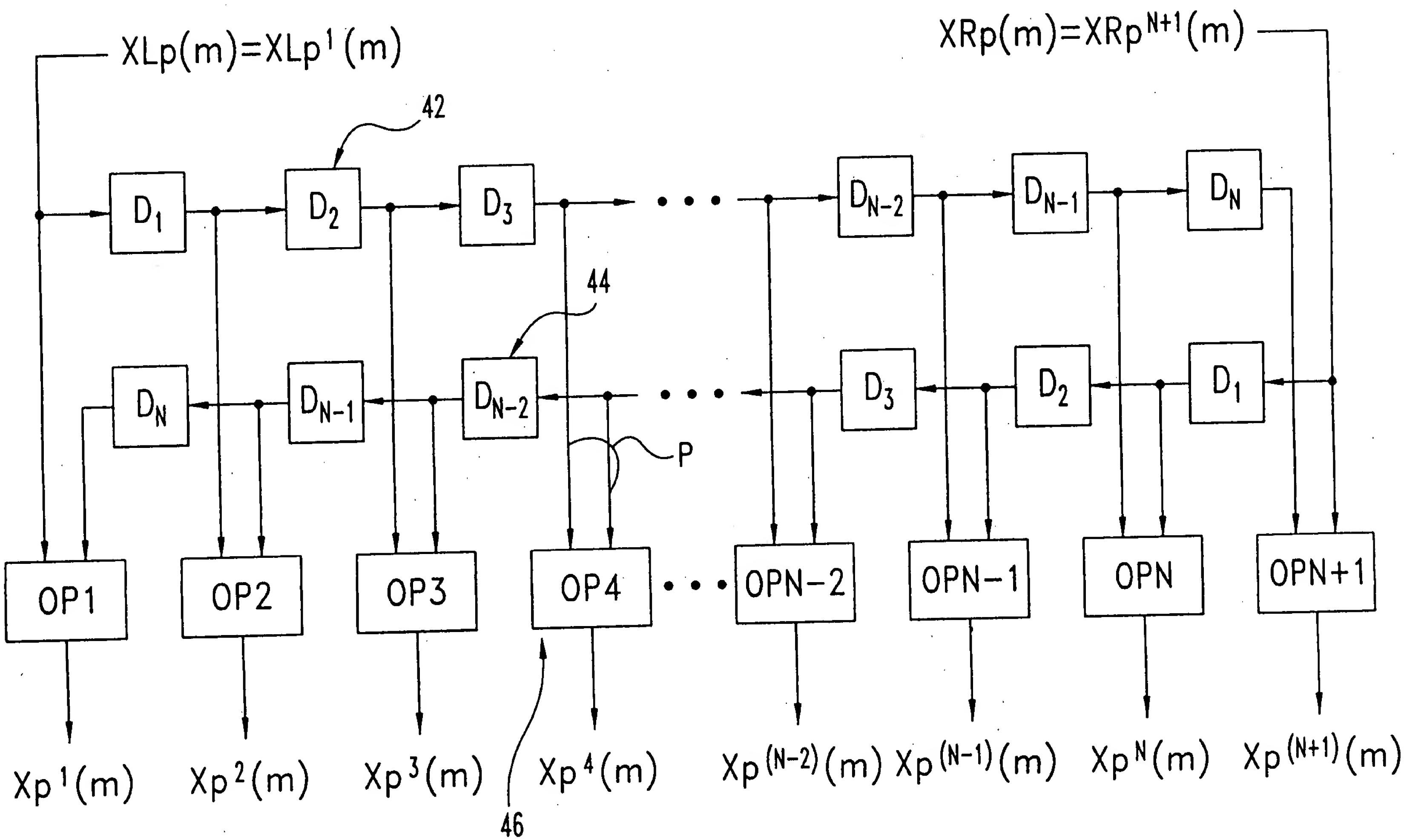

22
24
L
R
Filter 32a
Filter 32b
A/D 34a
A/D 34b
DFT 36a
DFT 36b
Dual Delay Line 40
Localization 70
Switch 80
D/A 84
Inverse DFT 82
30
90
AMP 92
Speaker 94

Fig. 2



XLp(m)=XLp¹(m)
XRp(m)=XRpᴺ⁺¹(m)
42
44
46
40
P
D1
D2
D3
DN-2
DN-1
DN
OP1
OP2
OP3
OP4
OPN-2
OPN-1
OPN
OPN+1
Xp¹(m)
Xp²(m)
Xp³(m)
Xp⁴(m)
Xp(N-2)(m)
Xp(N-1)(m)
XpN(m)
Xp(N+1)(m)

Fig. 3

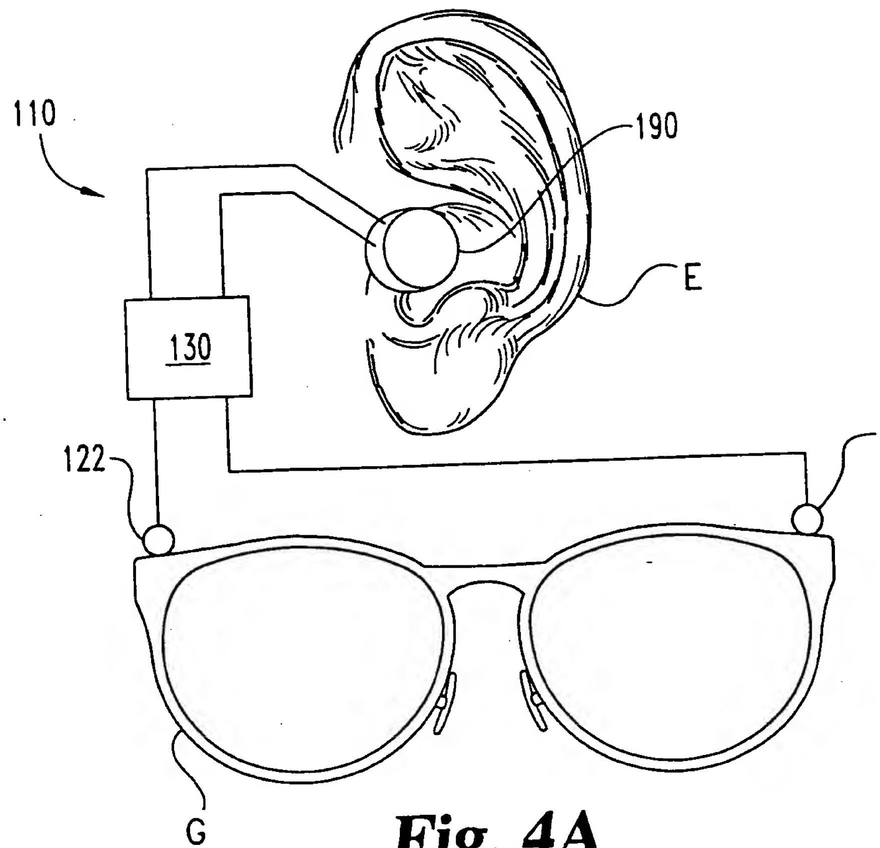


Fig. 4A

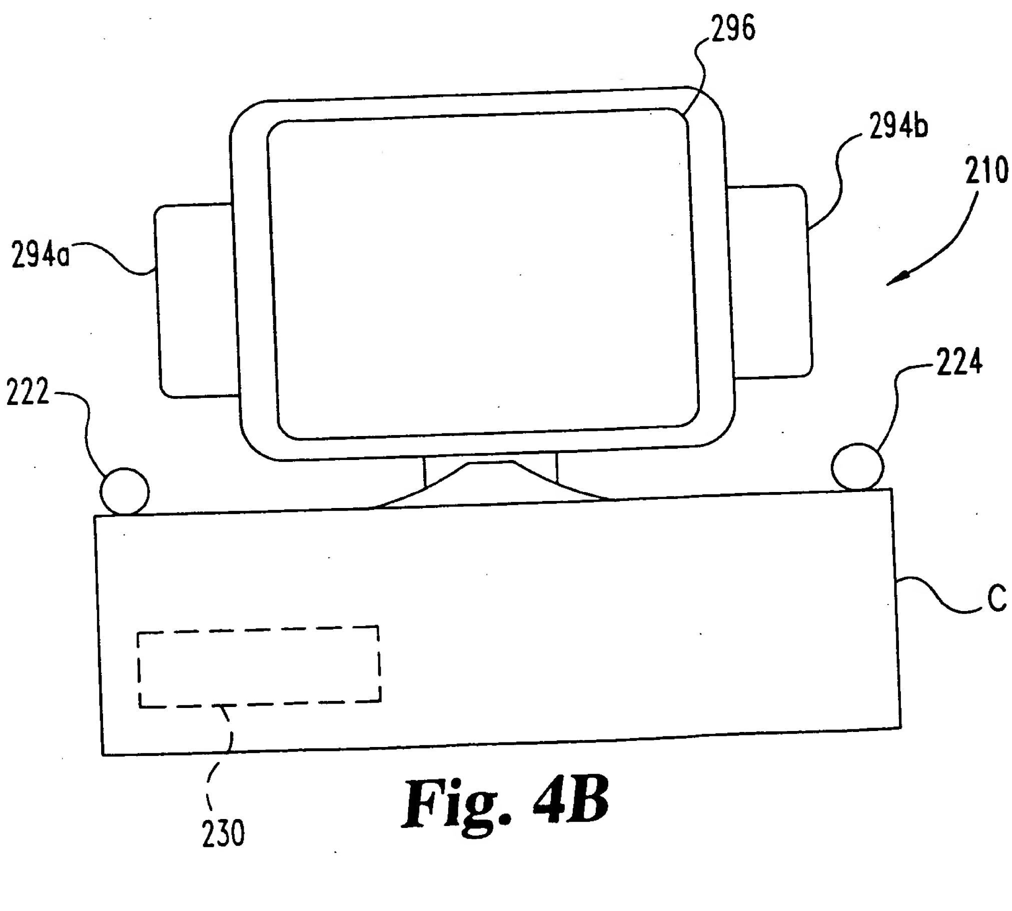


Fig. 4B

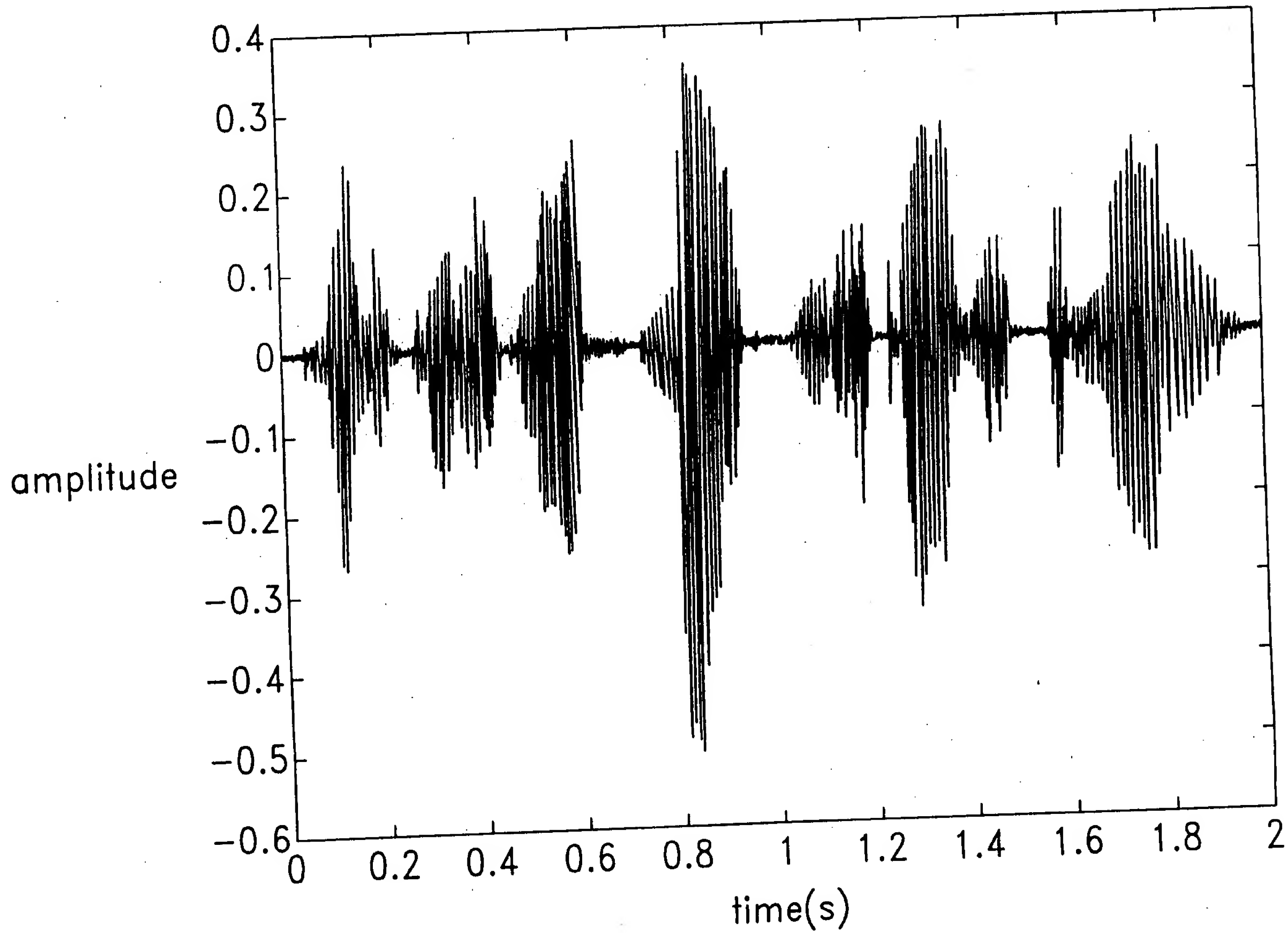

0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

Fig. 5

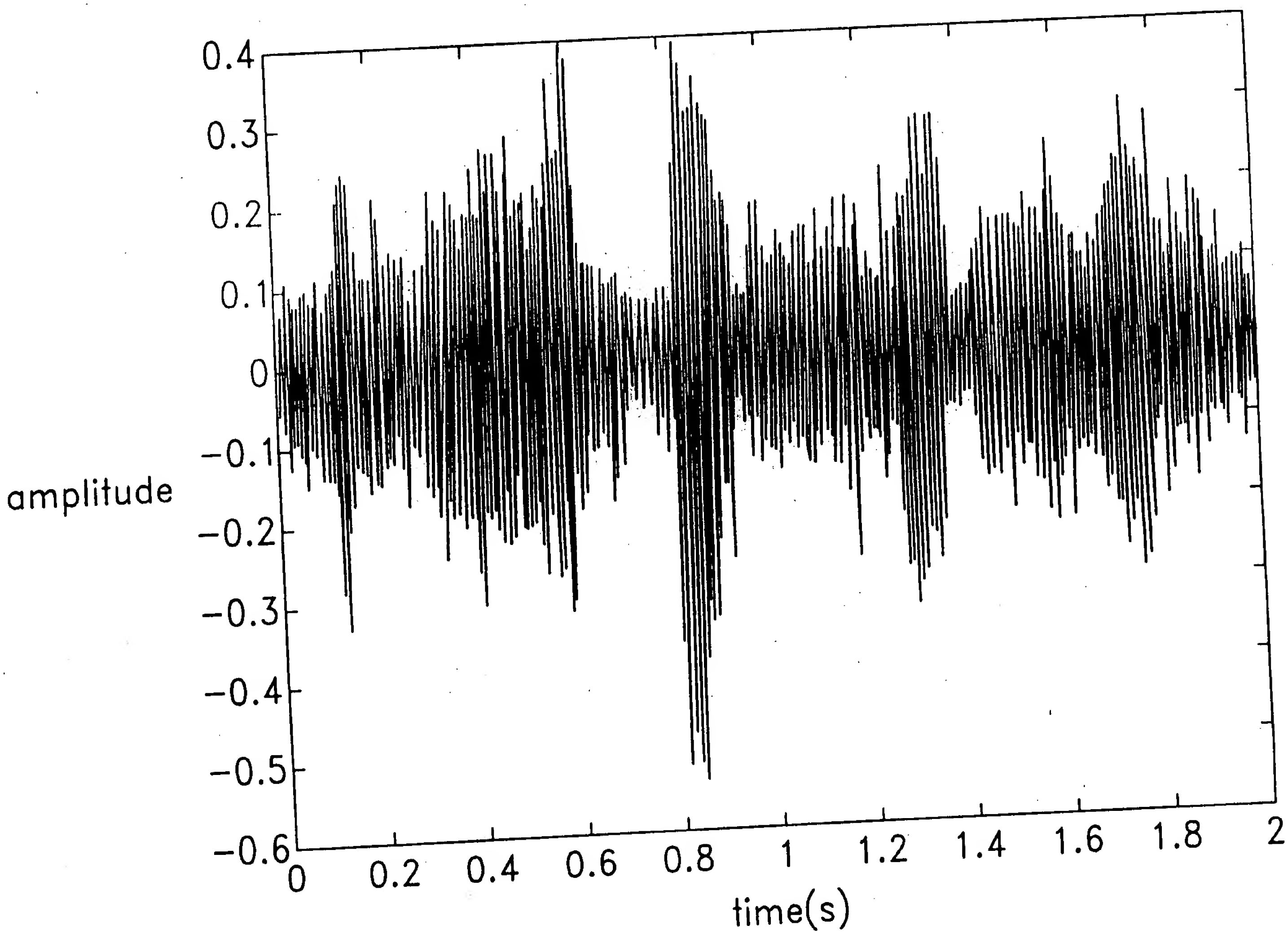
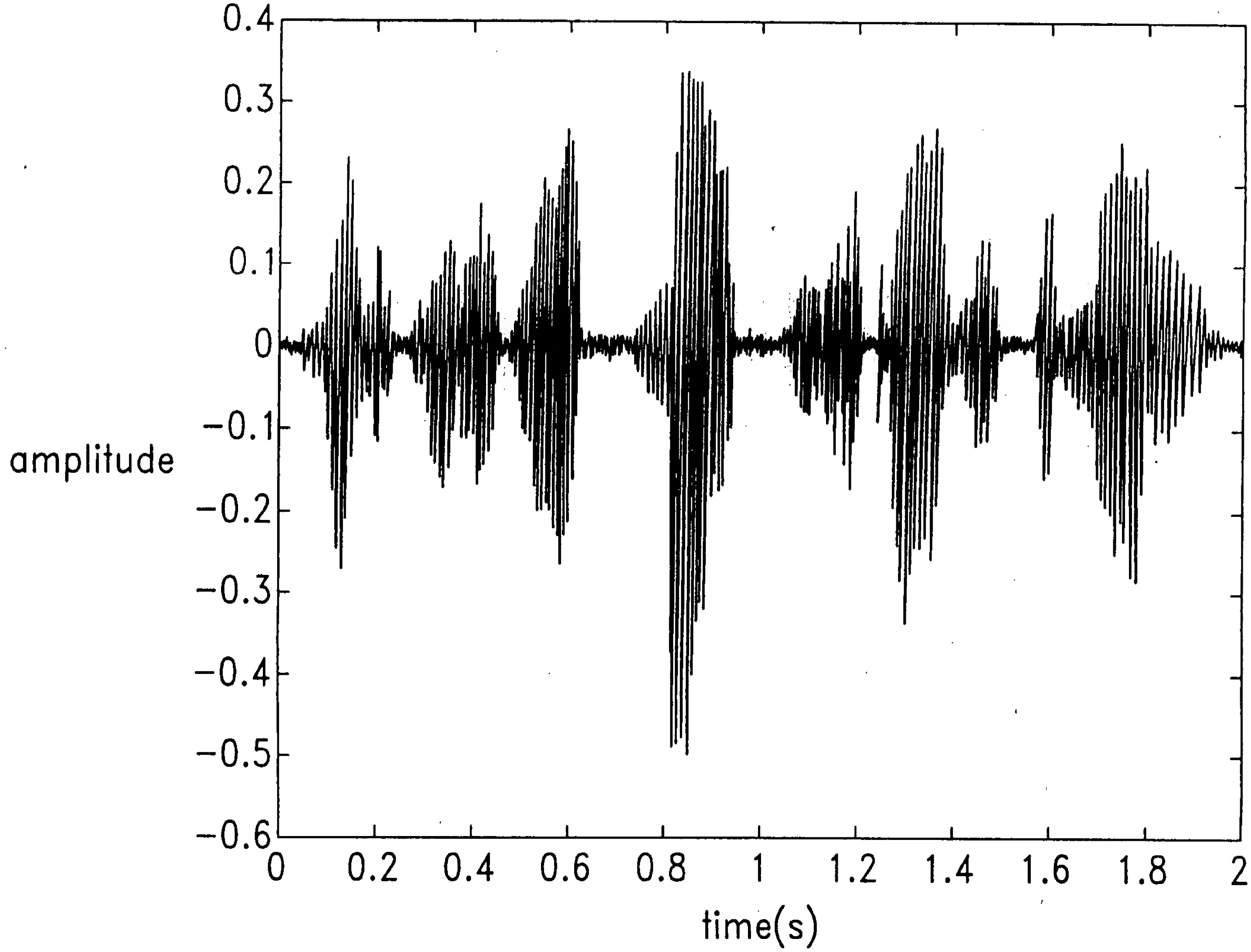

0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

Fig. 6



0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

Fig. 7

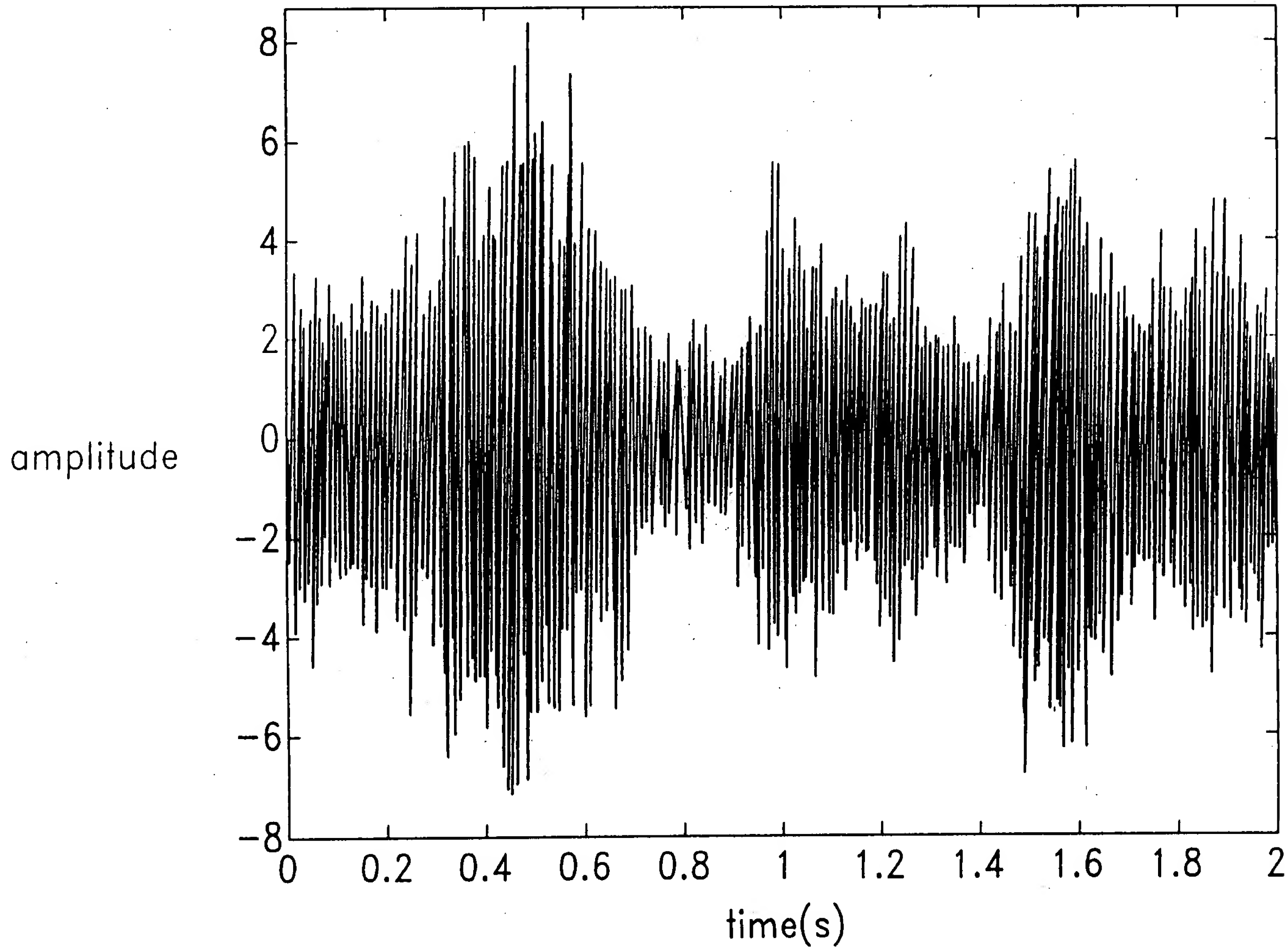


Fig. 8

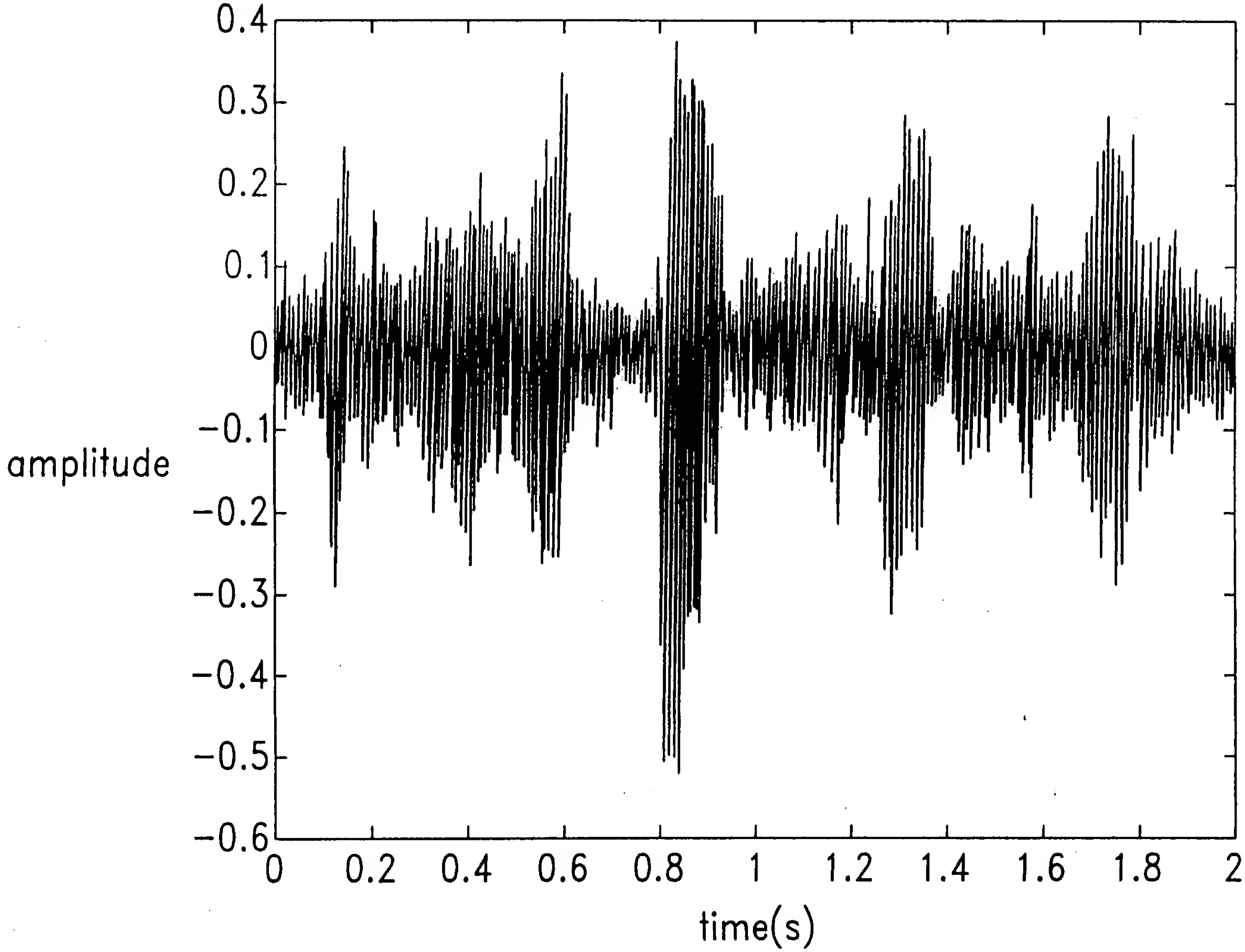
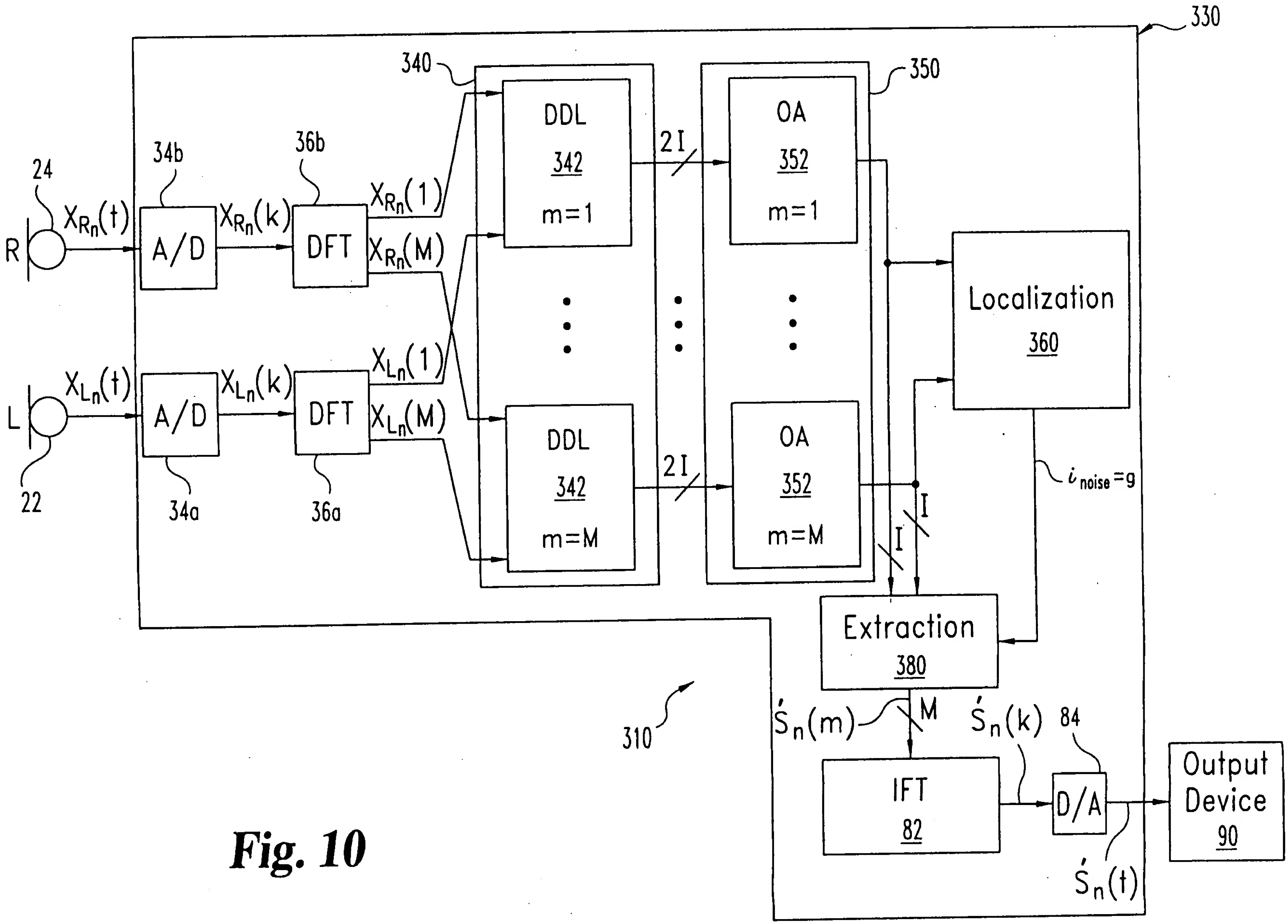
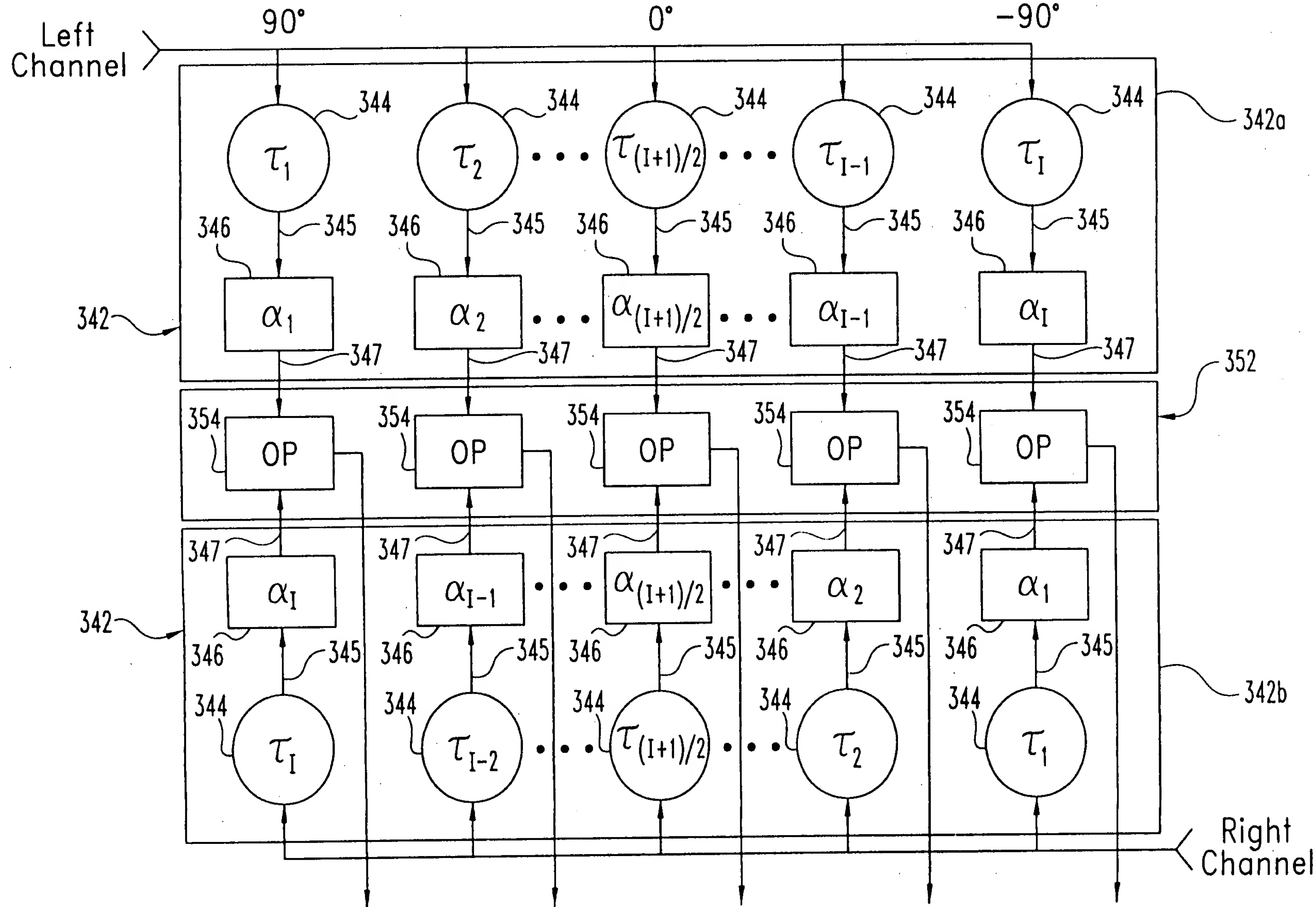

0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

Fig. 9



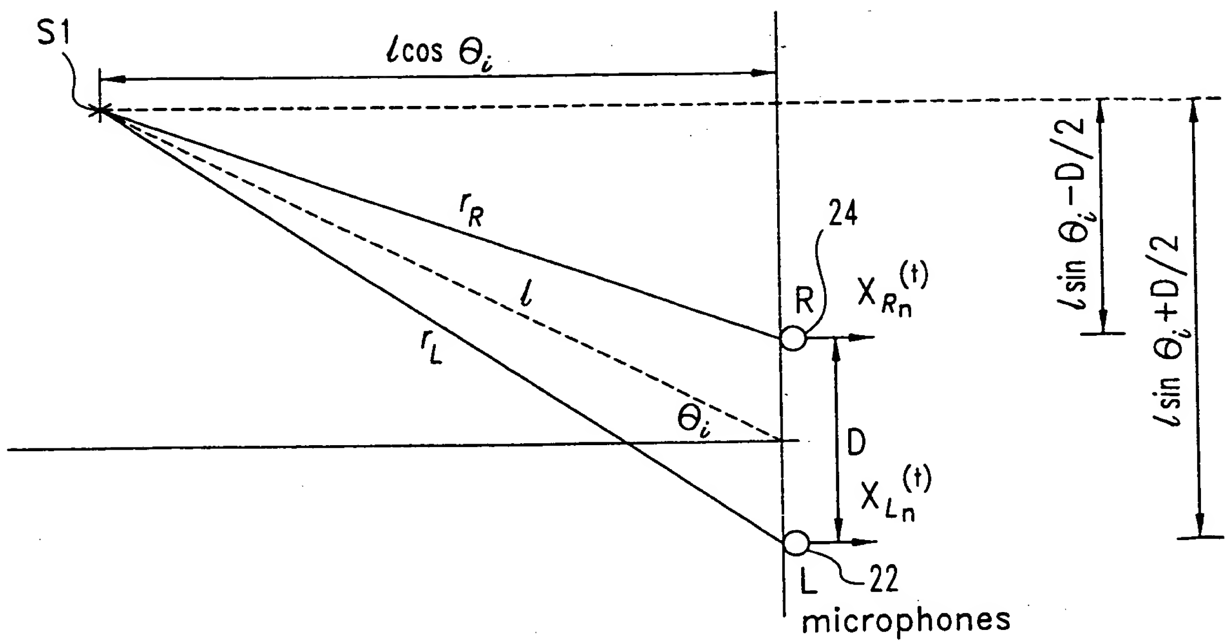
330
340
350
24
R
L
22
$X_{R_n}(t)$
$X_{L_n}(t)$
34b
34a
A/D
A/D
$X_{R_n}(k)$
$X_{L_n}(k)$
36b
36a
DFT
DFT
$X_{R_n}(1)$
$X_{R_n}(M)$
$X_{L_n}(1)$
$X_{L_n}(M)$
DDL
342
m=1
DDL
342
m=M
2I
2I
OA
352
m=1
OA
352
m=M
Localization
360
i_{noise}=g
I
I
Extraction
380
$\acute{S}_n(m)$
M
IFT
82
$\acute{S}_n(k)$
84
D/A
$\acute{S}_n(t)$
Output
Device
90
310

Fig. 10



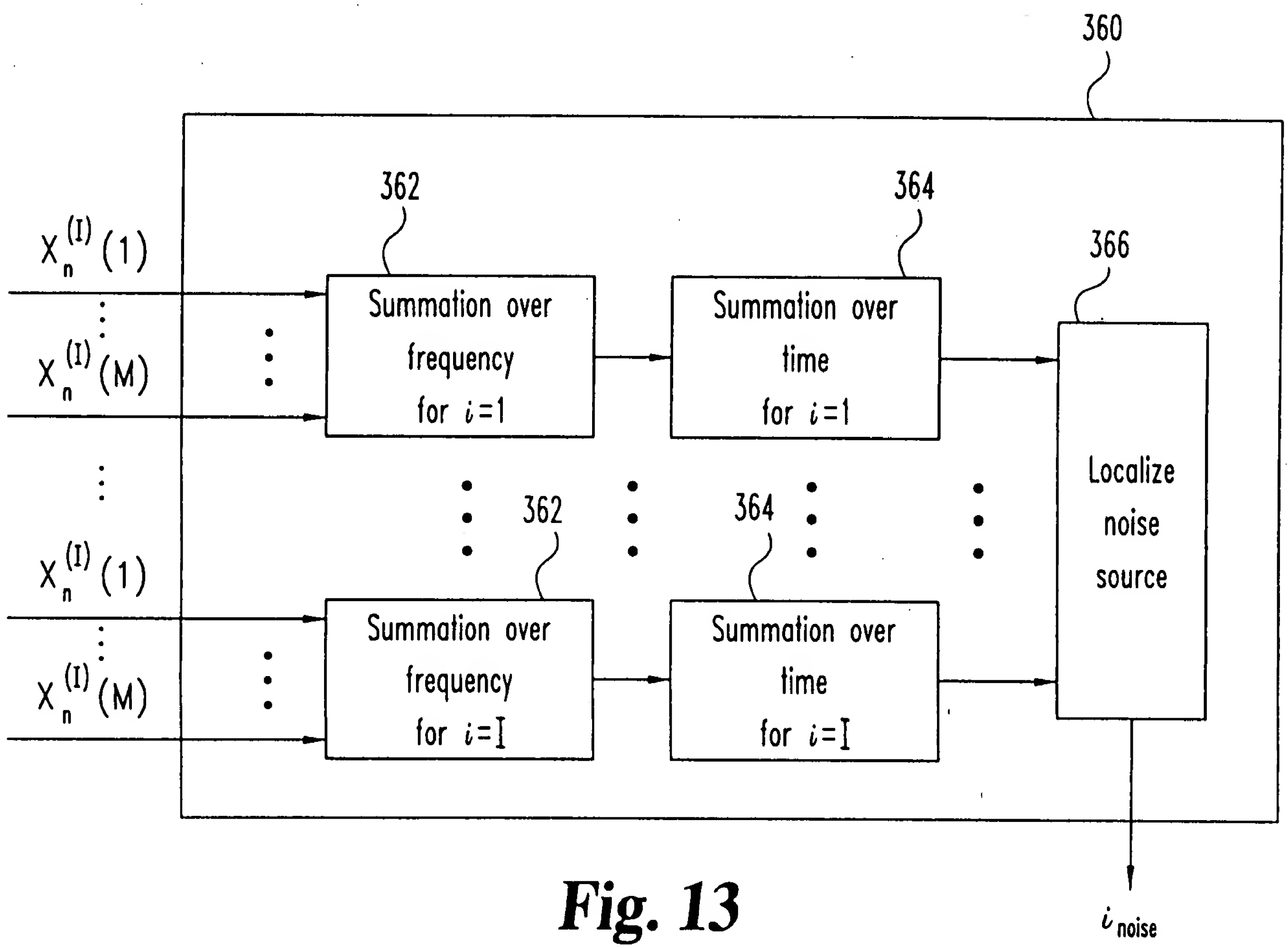
Left Channel
90°
0°
-90°
344
345
346
347
342a
342
352
354
OP
342b
Right Channel

Fig. 11



S1
$l\cos\Theta_i$
r_R
l
r_L
Θ_i
24
R
$X_{R_n}^{(t)}$
D
$X_{L_n}^{(t)}$
L
22
microphones
$l\sin\Theta_i - D/2$
$l\sin\Theta_i + D/2$

Fig. 12



360
362
364
366
$X_n^{(I)}(1)$
$X_n^{(I)}(M)$
$X_n^{(I)}(1)$
$X_n^{(I)}(M)$
Summation over frequency for $i=1$
Summation over time for $i=1$
362
364
Summation over frequency for $i=I$
Summation over time for $i=I$
Localize noise source
i_{noise}

Fig. 13

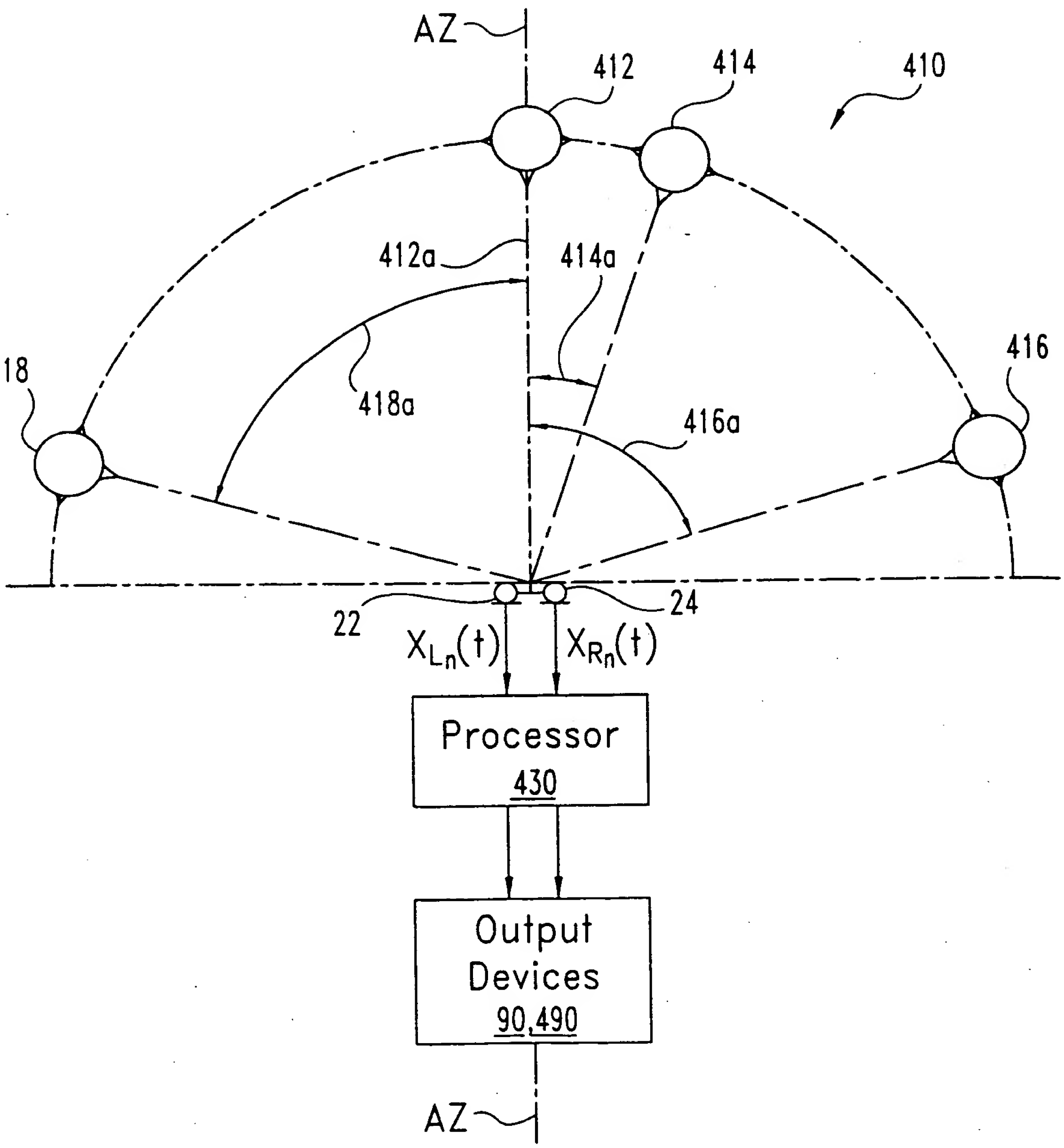
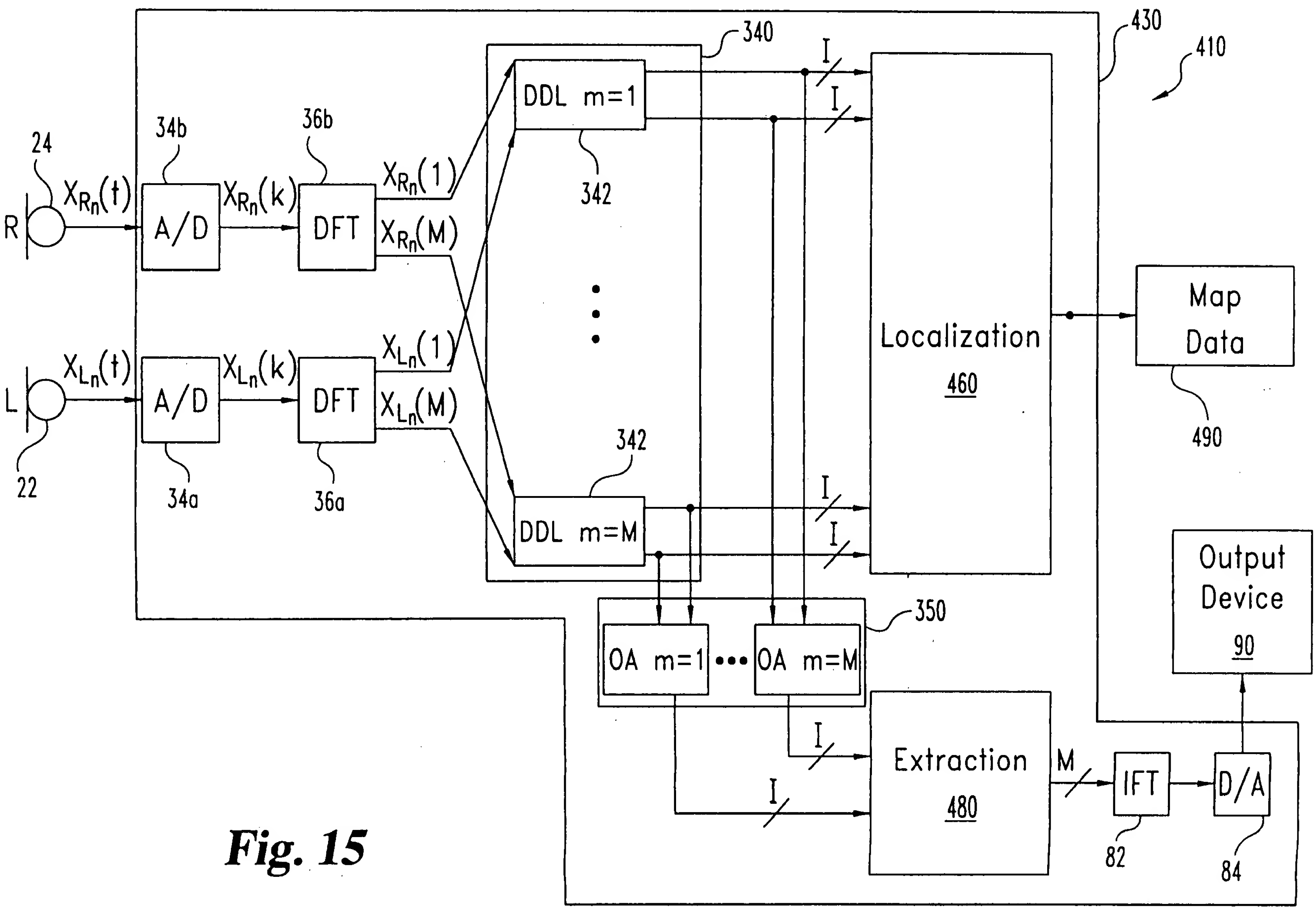
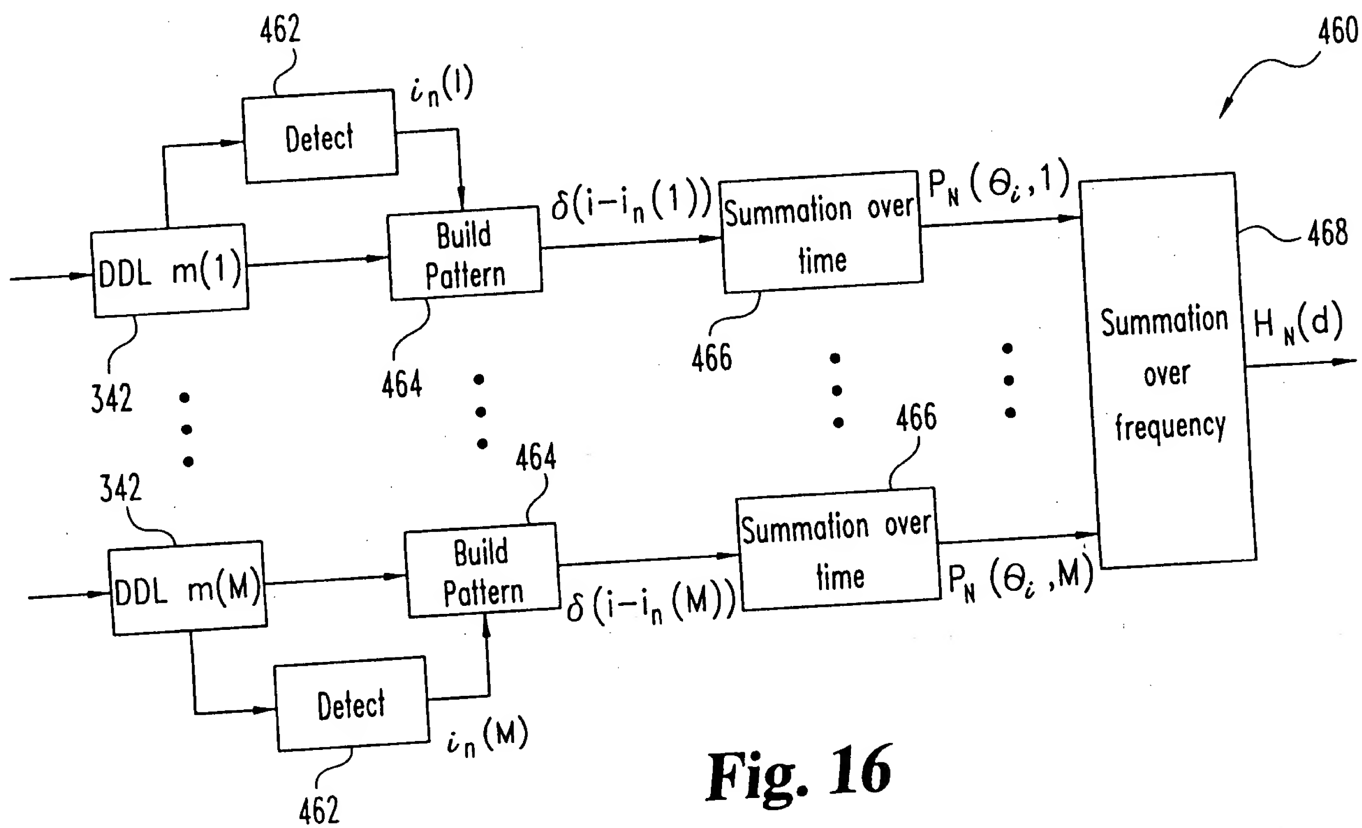

AZ
412
414
410
412a
414a
418
416
418a
416a
22
24
$X_{L_n}(t)$
$X_{R_n}(t)$
Processor
430
Output
Devices
90,490
AZ

Fig. 14



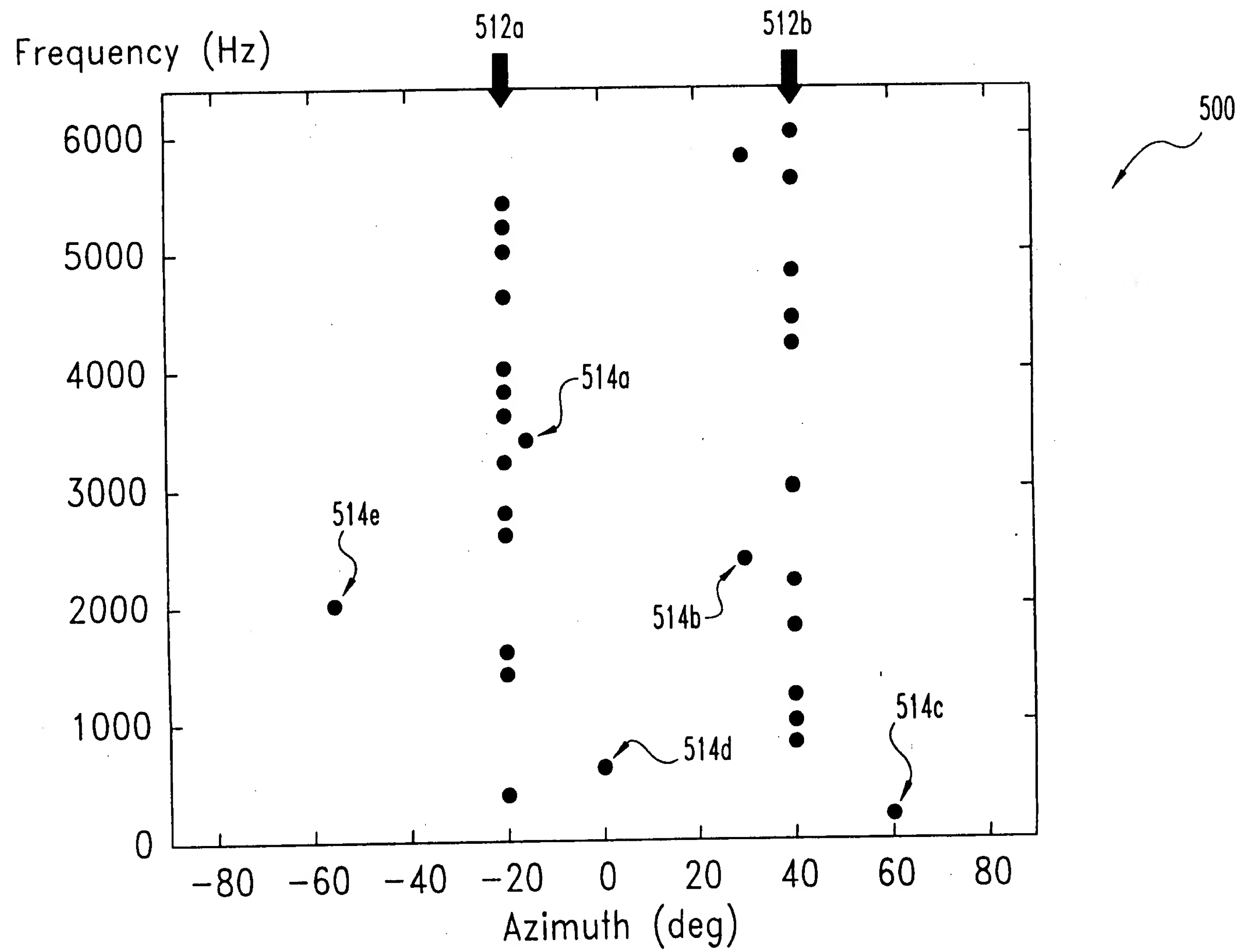
24
R
22
L
34b
34a
36b
36a
A/D
A/D
DFT
DFT
340
DDL m=1
DDL m=M
342
342
I
Localization
460
Map
Data
490
430
410
350
OA m=1
OA m=M
Extraction
480
M
IFT
D/A
82
84
Output
Device
90

Fig. 15



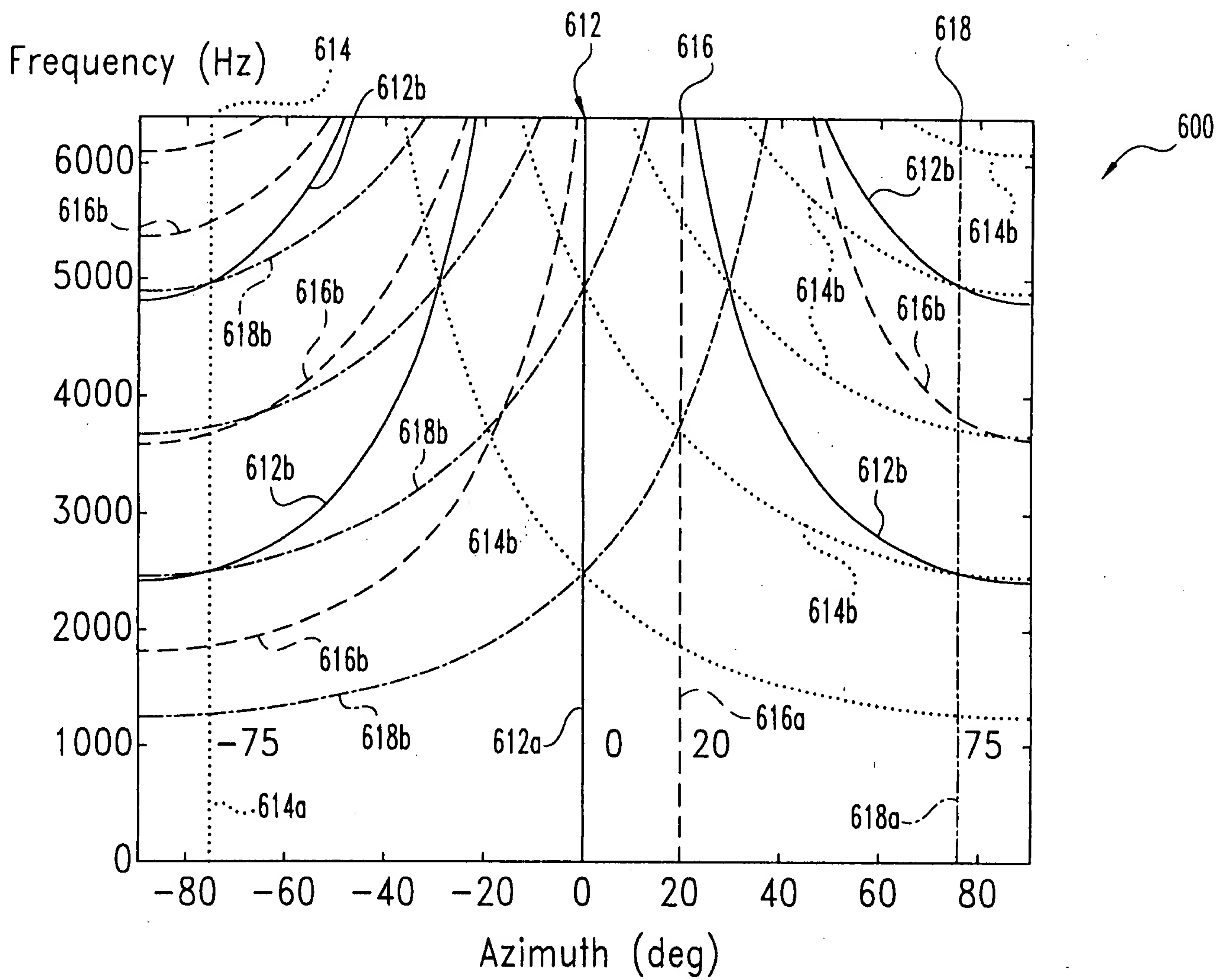
462
Detect
$i_n(l)$
460
DDL m(1)
Build Pattern
$\delta(i-i_n(1))$
Summation over time
$P_N(\Theta_i,1)$
468
Summation over frequency
$H_N(d)$
342
464
466
342
464
466
DDL m(M)
Build Pattern
$\delta(i-i_n(M))$
Summation over time
$P_N(\Theta_i,M)$
Detect
$i_n(M)$
462

Fig. 16



Frequency (Hz)
512a
512b
500
6000
5000
4000
3000
2000
1000
0
514a
514b
514c
514d
514e
-80
-60
-40
-20
0
20
40
60
80
Azimuth (deg)

Fig. 17



Frequency (Hz)
6000
5000
4000
3000
2000
1000
0
-80
-60
-40
-20
0
20
40
60
80
Azimuth (deg)
600
612
614
616
618
612a
614a
616a
618a
612b
614b
616b
618b
-75
0
20
75

Fig. 18

Test	Desired Source (°)	Intelligibility-Weighted Signal Cancellation (dB)				Intelligibility-Weighted Noise Cancellation (dB)	Net Intelligibility-Weighted Gain (dB)
		M1 "armchair"	M2 "playground"	F1 "pancake"	F2 "woodwork"		
#1		-75	0	20	75		
	-75	0.22	5.27	5.43	5.19	8.09	7.86
	0	7.94	-0.00	5.39	3.61	9.40	9.40
	20	8.24	3.63	-0.02	4.27	9.03	9.05
	75	8.37	4.69	5.17	0.05	10.16	10.11
#2		30	-45	60	-10		
	30	0.01	5.56	4.62	5.88	8.25	8.24
	-45	10.43	0.04	5.67	5.63	10.31	10.27
	60	11.24	5.56	0.06	5.70	10.65	10.59
	-10	9.32	4.29	3.51	-0.06	9.52	9.59
#3		10	-80	-50	45		
	10	0.01	4.88	5.25	3.05	7.79	7.78
	-80	10.00	0.15	5.53	3.29	10.66	10.52
	-50	10.10	2.99	0.03	3.39	9.54	9.51
	45	10.77	5.41	6.44	0.07	11.72	11.66
#4		-30	15	5	-60		
	-30	0.02	6.11	6.14	4.87	8.48	8.46
	15	9.55	-0.02	5.25	4.40	10.22	10.24
	5	9.24	2.99	-0.01	3.95	9.65	9.66
	-60	9.97	5.68	6.73	0.03	11.16	11.13
#5		-25	25	-70	80		
	-25	0.02	5.86	4.78	4.73	8.09	8.07
	25	9.07	-0.01	4.98	3.51	9.39	9.40
	-70	10.09	4.66	0.08	4.31	9.99	9.91
	80	9.40	4.90	4.61	0.02	10.27	10.25

TABLE I

Fig. 19

Test	Desired Source (°)	Intelligibility-Weighted Signal Cancellation (dB) M1 "armchair"	M2 "playground"	F1 "pancake"	F2 "woodwork"	Intelligibility-Weighted Noise Cancellation (dB)	Net Intelligibility-Weighted Gain (dB)
#1		-75	0	20	75		
	-75	0.13	5.36	5.91	5.93	8.45	8.32
	0	8.10	0.00	5.20	4.09	9.71	9.71
	20	8.28	3.43	-0.01	4.76	9.16	9.18
	75	8.36	4.49	5.15	0.04	10.32	10.28
#2		30	-45	60	-10		
	30	0.01	5.50	4.61	5.79	8.22	8.20
	-45	10.48	0.04	5.67	5.54	10.31	10.28
	60	11.21	5.56	0.07	5.63	10.66	10.59
	-10	9.28	4.25	3.44	-0.07	9.53	9.60
#3		10	-80	-50	45		
	10	0.05	5.06	4.83	3.17	7.90	7.85
	-80	9.82	0.10	5.59	3.63	10.85	10.76
	-50	9.48	3.60	0.06	3.31	10.15	10.09
	45	10.51	5.47	6.38	0.05	11.95	11.90
#4		-30	15	5	-60		
	-30	0.02	6.11	6.13	5.11	8.56	8.53
	15	9.53	-0.00	5.31	4.33	10.25	10.25
	5	9.18	2.95	-0.00	4.01	9.55	9.56
	-60	9.70	5.33	6.07	0.01	10.92	10.91
#5		-25	25	-70	80		
	-25	0.01	5.82	4.96	5.37	8.24	8.24
	25	8.77	-0.00	5.19	4.29	9.43	9.43
	-70	9.77	4.85	0.05	4.74	10.18	10.13
	80	9.02	4.58	4.73	-0.00	10.38	10.39

TABLE II

Fig. 20

Test	Desired Source (°)	Intelligibility-Weighted Signal Cancellation (dB)				Intelligibility-Weighted Noise Cancellation (dB)	Net Intelligibility Weighted Gain (dB)
		F3 "*stairway*"	F4 "*mushroom*"	M3 "*birthday*"	M4 "*sidewalk*"		
#1		-75	0	20	75		
	-75	0.09	8.47	7.85	6.17	9.51	9.42
	0	5.69	-0.01	6.45	5.38	8.30	8.31
	20	6.12	7.33	-0.03	5.09	9.65	9.69
	75	7.34	8.62	8.24	0.08	11.03	10.95
#2		30	-45	60	-10		
	30	0.01	8.04	6.31	6.63	9.24	9.22
	-45	7.48	0.13	6.05	6.06	9.02	8.89
	60	7.78	8.52	0.10	6.85	10.51	10.41
	-10	7.86	7.53	6.65	-0.03	10.02	10.05
#3		10	-80	-50	45		
	10	-0.11	5.96	6.32	6.99	7.48	7.59
	-80	6.06	0.11	5.86	6.74	9.11	9.00
	-50	6.71	4.33	0.06	7.18	7.95	7.88
	45	7.07	6.42	6.78	0.05	8.98	8.93
#4		-30	15	5	-60		
	-30	0.02	8.37	8.24	5.70	9.33	9.31
	15	6.56	0.02	6.51	5.81	8.87	8.85
	5	6.30	5.27	0.01	5.68	8.77	8.76
	-60	7.61	8.22	8.41	0.05	10.64	10.58
#5		-25	25	-70	80		
	-25	0.00	7.68	6.67	6.50	9.75	9.74
	25	6.14	-0.03	6.03	4.46	8.20	8.22
	-70	5.60	6.54	0.22	4.58	9.05	8.83
	80	6.85	7.42	6.12	0.08	9.91	9.84

TABLE III

Fig. 21

Test	Desired Source (°)	Intelligibility-Weighted Signal Cancellation (dB)				Intelligibility-Weighted Noise Cancellation (dB)	Net Intelligibility-Weighted Gain (dB)
		F3 "*stairway*"	F4 "*mushroom*"	M3 "*birthday*"	M4 "*sidewalk*"		
#1		-75	0	20	75		
	-75	0.11	6.41	6.78	7.15	8.17	8.06
	0	5.68	0.00	5.12	6.26	8.30	8.30
	20	6.10	5.82	-0.03	6.17	8.39	8.42
	75	7.40	6.04	6.60	0.09	8.44	8.35
#2		30	-45	60	-10		
	30	0.02	8.06	6.07	6.41	9.16	9.14
	-45	7.55	0.11	5.63	6.65	9.05	8.95
	60	7.47	8.48	0.08	6.64	10.55	10.47
	-10	7.57	7.60	6.31	-0.04	10.04	10.07
#3		10	-80	-50	45		
	10	-0.11	5.94	3.79	6.87	6.41	6.52
	-80	6.30	0.10	4.04	6.56	8.07	7.97
	-50	6.69	4.68	0.07	6.89	8.29	8.22
	45	7.16	6.07	4.94	0.03	8.02	7.99
#4		-30	15	5	-60		
	-30	0.02	8.45	7.08	6.18	9.10	9.08
	15	7.18	0.00	5.07	6.31	8.02	8.02
	5	6.27	5.21	0.03	5.47	8.66	8.64
	-60	7.89	8.34	7.19	0.05	10.35	10.30
#5		-25	25	-70	80		
	-25	0.01	7.79	6.63	6.57	9.67	9.66
	25	6.08	-0.03	5.90	4.82	8.25	8.28
	-70	5.58	6.66	0.15	4.43	9.18	9.03
	80	6.67	7.63	5.87	0.07	9.81	9.74

TABLE IV

Fig. 22

BINAURAL SIGNAL PROCESSING SYSTEM AND METHOD

BACKGROUND OF THE INVENTION

The present invention is directed to the processing of acoustic signals, and more particularly, but not exclusively, relates to the separation of acoustic signals emanating from different sources by detecting a mixture of the acoustic signals at multiple locations.

The difficulty of extracting a desired signal in the presence of interfering signals is a long-standing problem confronted by acoustic engineers. This problem impacts the design and construction of many kinds of devices such as systems for voice recognition and intelligence gathering. Especially troublesome is the separation of desired sound from unwanted sound with hearing aid devices. Generally, hearing aid devices do not permit selective amplification of a desired sound when contaminated by noise from a nearby source -- particularly when the noise is more intense. This problem is even more severe when the desired sound is a speech signal and the nearby noise is also the result of speech (e.g. babble). As used herein, "noise" refers not only to random or non deterministic signals, but also to undesired signals and signals interfering with the perception of a desired signal.

One attempted solution to this problem has been the application of a single, highly directional microphone to enhance directionality of the hearing aid receiver. This approach has only a very limited capability. As a result, spectral subtraction, comb filtering, and speech-production modeling have been explored to enhance single microphone performance. Nonetheless, these approaches still generally fail to improve intelligibility of a desired speech signal, particularly when the signal and noise source are in close proximity.

Another approach has been to arrange a number of microphones in a selected spatial relationship to form a type of directional detection beam. Unfortunately, when limited to a

size practical for hearing aids, beam forming arrays also have limited capacity to separate signals which are close together -- especially if the noise is more intense than a desired speech signal. In addition, in the case of one noise source in a less reverberant environment, the noise cancellation provided by the beam-former varies with the location of the noise source in relation to the microphone array. R.W. Stadler and W.M. Rabinowitz, On the Potential of Fixed Arrays for Hearing Aids, 94 Journal Acoustical Society of America 1332 (September 1993), and W. Soede et al., Development of a Directional Hearing Instrument Based on Array Technology, 94 Journal of Acoustical Society of America 785 (August 1993) are cited as additional background concerning the beam forming approach.

Still another approach has been the application of two microphones displaced from each other to provide two signals to emulate certain aspects of the binaural hearing system common to humans and many types of animals. Although certain aspects of biologic binaural hearing are still not fully understood, it is believed that the ability to localize sound sources is based on evaluation of binaural time delays and sound levels across different frequency bands associated with each of the two sound signals. The localization of sound sources with systems based on these interaural time and intensity differences is discussed in W. Lindemann, Extension of a Binaural Cross-Correlation Model by Contralateral Inhibition - I. Simulation of Lateralization for Stationary Signals, 80 Journal of the Acoustical Society of America 1608 (December 1986). Nonetheless, the separation of a desired signal from noise or interfering sound still presents a significant problem once the sound sources are localized.

For example, the system set forth in Markus Bodden, Modeling Human Sound-Source Localization and the Cocktail-Party-Effect, 1 Acta Acustica 43 (February/April 1993) employs a Wiener filter including a windowing process in an attempt to derive a desired signal from binaural input signals once the location of the desired signal has been established. Unfortunately, this approach results in significant deterioration of desired speech fidelity. Also, the system has only been demonstrated to suppress noise of equal intensity to the

desired signal at an azimuthal separation of at least 30 degrees. A more intense noise emanating from a source spaced closer than 30 degrees from the desired source still appears to present a problem. Moreover, the proposed algorithm of the Bodden system is computationally intense -- posing a serious question of whether it can be practically embodied in a hearing aid device.

Another example of a two microphone system is found in D. Banks, Localisation and Separation of Simultaneous Voices with Two Microphones, IEE Proceedings-I, 140 (1993). This system employs a windowing technique to estimate the location of a sound source when there are non overlapping gaps in its spectrum compared to the spectrum of interfering noise. This system cannot perform localization when wide-band signals lacking such gaps are involved. In addition, the Banks article fails to provide details of the algorithm for reconstructing the desired signal. U.S. Patent Nos. 5,479,522 to Lindemann et al.; 5,325,436 to Soli et al.; 5,289,544 to Franklin; and 4,773,095 to Zwicker et al. are cited as sources of additional background concerning dual microphone hearing aid systems.

These binaural systems still fail to provide for the extraction of an intelligible speech signal subject to acoustic interference emanating from a nearby noise source. Thus, a need remains for a way to extract a desired acoustic signal from a noisy environment which minimizes degradation of the desired signal fidelity and which may be practically embodied into a device such as a hearing aid.

SUMMARY OF THE INVENTION

One feature of the present invention is utilizing two sensors to provide corresponding binaural signals from which the relative separation of a first acoustic source from a second acoustic source may be established as a function of time, and the spectral content of a desired acoustic signal from the first source may be representatively extracted. One aspect of this feature is that the desired acoustic signal may be successfully extracted even if a nearby noise source is of greater relative intensity.

Another feature of the present invention is detecting an acoustic excitation at a first location to provide a corresponding first signal and at a second location to provide a corresponding second signal. This excitation includes a desired acoustic signal from a first source and an interfering acoustic signal from a second source spaced apart from the first source. The second source is localized relative to the first source as a function of the first and second signals. A characteristic signal is generated which is representative of the desired acoustic signal during the localization.

Still another feature is delaying the first and second signals by a number of time intervals to correspondingly establish a number of delayed first signals and a number of delayed second signals. A time increment corresponding to the separation of the first and second sources is determined by comparing the delayed first signals to the delayed second signals. An output signal representative of the desired signal is generated as a function of the time increment. Furthermore, a signal pair indicative of the location of the second source may be selected that has a first member selected from the delayed first signals and a second member from the delayed second signals. The output signal may be generated as a function of this signal pair.

In yet another feature, a processing system utilizes a first and second sensor at different locations to provide a binaural representation of an acoustic signal which includes a desired signal emanating from a selected source and an interfering signal emanating from a

interfering source. A processor generates a discrete first spectral signal and a discrete second spectral signal from the sensor signals. The processor delays the first and second spectral signals by a number of time intervals to generate a number of delayed first signals and a number of delayed second signals and provide a time increment signal. The time increment signal corresponds to separation of the selected source from the noise source. The processor generates an output signal as a function of the time increment signal, and an output device responds to the output signal to provide a sensory output representative of the desired signal.

Among the other features of the present invention is a system to position a first and second sensor relative to a first signal source with the first and second sensor being spaced apart from each other and a second signal source being spaced apart from the first signal source. A first signal is provided from the first sensor and a second signal is provided from the second sensor. The first and second signals each represent a composite acoustic signal including a desired signal from the first signal source and an unwanted signal from the second signal source. A number of spectral signals are established from the first and second signals as a function of a number of frequencies. Each of the spectral signals, such as those corresponding to outputs of a delay line, represent a different position relative to the first signal source. A member of the spectral signals representative of position of the second signal source is determined, and an output signal is generated from the member which is representative of the first signal. This feature facilitates extraction of a desired signal from a spectral signal determined as part of the localization of the interfering source. As a result, localization calculations constitute the bulk of the signal processing because, once localization of the interfering source is performed, the desired signal is estimated directly from one of the intermediate localization operands. This approach avoids the extensive post-localization computations required by many binaural systems.

Accordingly, it is one object of the present invention to provide for the extraction of a desired acoustic signal from a noisy environment.

Another object is to provide a device for the separation of acoustic signals by detecting a combination of these signals at two locations. This device may be used to aid impaired hearing.

Further objects, features, and advantages of the present invention shall become apparent from the detailed drawings and descriptions provided herein.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagrammatic view of a first embodiment of the present invention.

FIG. 2 is a signal flow diagram of an extraction process performed by the embodiment of FIG. 1.

FIG. 3 is schematic representation of the dual delay line of FIG. 2.

FIGS. 4A and 4B depict other embodiments of the present invention corresponding to hearing aid and computer voice recognition applications, respectively.

FIG. 5 is a graph of a speech signal in the form of a sentence about 2 seconds long.

FIG. 6 is a graph of a composite signal including babble noise and the speech signal of FIG. 5 at a 0 dB signal-to-noise ratio with the babble noise source at about a 60 azimuth relative to the speech signal source.

FIG. 7 is a graph of a signal representative of the speech signal of FIG. 5 after extraction from the composite signal of FIG. 6.

FIG. 8 is a graph of a composite signal including babble noise and the speech signal of FIG. 5 at a -30 dB signal-to-noise ratio with the babble noise source at a 2 degree azimuth relative to the speech signal source.

FIG. 9 is a graphic depiction of a signal representative of the sample speech signal of FIG. 5 after extraction from the composite signal of FIG. 8.

DESCRIPTION OF THE PREFERRED EMBODIMENT

For the purposes of promoting an understanding of the principles of the invention, reference will now be made to the embodiment illustrated in the drawings and specific language will be used to describe the same. It will nevertheless be understood that no limitation of the scope of the invention is thereby intended. Any alterations and further modifications in the described device, and any further applications of the principles of the invention as described herein are contemplated as would normally occur to one skilled in the art to which the invention relates.

Fig. 1 illustrates an acoustic signal processing system 10 of the present invention. System 10 is configured to extract a desired acoustic signal from source 12 despite interference or noise emanating from nearby source 14. System 10 includes a pair of acoustic sensors 22, 24 configured to detect acoustic excitation that includes signals from sources 12, 14. Sensors 22, 24 are operatively coupled to processor 30 to process signals received therefrom. Also, processor 30 is operatively coupled to output device 90 to provide a signal representative of a desired signal from source 12 with reduced interference from source 14 as compared to composite acoustic signals presented to sensors 22, 24 from sources 12, 14.

Sensors 22, 24 are spaced apart from one another by distance D along lateral axis T. Midpoint M represents the half way point along distance D from sensor 22 to sensor 24. Reference axis R1 is aligned with source 12 and intersects axis T perpendicularly through midpoint M. Axis N is aligned with source 14 and also intersects midpoint M. Axis N is positioned to form angle A with reference axis R1. Fig. 1 depicts an angle A of about 20 degrees. Notably, reference axis R1 may be selected to define a reference azimuthal position of zero degrees in an azimuthal plane intersecting sources 12, 14; sensors 22, 24; and containing axes T, N, R1. As a result,

source 12 is "on-axis" and source 14, as aligned with axis N, is "off-axis." Source 14 is illustrated at about a 20 degree azimuth relative to source 12.

Preferably sensors 22, 24 are fixed relative to each other and configured to move in tandem to selectively position reference axis R1 relative to a desired acoustic signal source. It is also preferred that sensors 22, 24 be a microphones of a conventional variety, such as omnidirectional dynamic microphones. In other embodiments, a different sensor type may be utilized as would occur to one skilled in the art.

Referring additionally to FIG. 2, a signal flow diagram illustrates various processing stages for the embodiment shown in FIG. 1. Sensors 22, 24 provide analog signals Lp(t) and Rp(t) corresponding to the left sensor 22, and right sensor 24, respectively. Signals Lp(t) and Rp(t) are initially input to processor 30 in separate processing channels L and R. For each channel L, R, signals Lp(t) and Rp(t) are conditioned and filtered in stages 32a, 32b to reduce aliasing, respectively. After filter stages 32a, 32b, the conditioned signals Lp(t), Rp(t) are input to corresponding Analog to Digital (A/D) converters 34a, 34b to provide discrete signals Lp(k), Rp(k), where k indexes discrete sampling events. In one embodiment, A/D stages 34a, 34b sample signals Lp(t) and Rp(t) at a rate of at least twice the frequency of the upper end of the audio frequency range to assure a high fidelity representation of the input signals.

Discrete signals Lp(k) and Rp(k) are transformed from the time domain to the frequency domain by a short-term Discrete Fourier Transform (DFT) algorithm in stages 36a, 36b to provide complex-valued signals XLp(m) and XRp(m). Signals XLp(m) and XRp(m) are evaluated in stages 36a, 36b at discrete frequencies f_m, where m is an index (m=1 to m=M) to discrete frequencies, and index p denotes the short-term spectral analysis time frame. Index p is arranged in reverse chronological order with the most recent time frame being p =1, the next most recent time frame being p = 2, and so forth. Preferably, frequencies M encompass the audible frequency range

and the number of samples employed in the short-term analysis is selected to strike an optimum balance between processing speed limitations and desired resolution of resulting output signals. In one embodiment, an audio range of 0.1 to 6 kHz is sampled in A/D stages 34a, 34b at a rate of at least 12.5 kHz with 512 samples per short-term spectral analysis time frame. In alternative embodiments, the frequency domain analysis may be provided by an analog filter bank employed before A/D stages 34a, 34b. It should be understood that the spectral signals XLp(m) and XRp(m) may be represented as arrays each having a 1xM dimension corresponding to the different frequencies f_m.

Spectral signals XLp(m) and XRp(m) are input to dual delay line 40 as further detailed in FIG. 3. FIG. 3 depicts two delay lines 42, 44 each having N number of delay stages. Each delay line 42, 44 is sequentially configured with delay stages D_1 through D_N. Delay lines 42, 44 are configured to delay corresponding input signals in opposing directions from one delay stage to the next, and generally correspond to the dual hearing channels associated with a natural binaural hearing process. Delay stages D_1, D_2, D_3, . . ., D_{N-2}, D_{N-1}, and D_N each delay an input signal by corresponding time delay increments τ_1, τ_2, τ_3, . . ., τ_{N-2}, τ_{N-1}, and τ_N, (collectively designated τ_i), where index i goes from left to right. For delay line 42, XLp(m) is alternatively designated $XLp^1(m)$. $XLp^1(m)$ is sequentially delayed by time delay increments τ_1, τ_2, τ_3, . . ., τ_{N-2}, τ_{N-1}, and τ_N to produce delayed outputs at the taps of delay line 42 which are respectively designated $XLp^2(m)$, $XLp^3(m)$, $Xlp^4(m)$, . . ., $XLp^{N-1}(m)$, $XLp^N(m)$, and $XLp^{N+1}(m)$; and collectively designated $XLp^i(m)$). For delay line 44, XRp(m) is alternatively designated $XRp^{N+1}(m)$. $XRp^{N+1}(m)$ is sequentially delayed by time delay increments τ_1, τ_2, τ_3, . . ., τ_{N-2}, τ_{N-1}, and τ_N to produce delayed outputs at the taps of delay line 44 which are respectively designated: $XRp^N(m)$, $XRp^{N-1}(m)$, $XRp^{N-2}(m)$, . . ., $XLp^3(m)$, $XLp^2(m)$, and $Xlp^1(m)$; and collectively designated $XRp^i(m)$. The input spectral

signals and the signals from delay line 42, 44 taps are arranged as input pairs to operation array 46. A pair of taps from delay lines 42, 44 is illustrated as input pair P in FIG. 3.

Operation array 46 has operation units (OP) numbered from 1 to N+1, depicted as OP1, OP2, OP3, OP4,..., OPN-2, OPN-1, OPN, OPN+1 and collectively designated operations OPi. Input pairs from delay lines 42, 44 correspond to the operations of array 46 as follows: $OP1[XLp^1(m), XRp^1(m)]$, $OP2[XLp^2(m), XRp^2(m)]$, $OP3[XLp^3(m), XRp^3(m)]$, $OP4[XLp^4(m), XRp^4(m)]$,..., $OPN\text{-}2[XLp^{(N-2)}(m), XRp^{(N-2)}(m)]$, $OPN\text{-}1[XLp^{(N-1)}(m), XRp^{(N-1)}(m)]$, $OPN[XLp^N(m), XRp^N(m)]$, and $OPN+1[XLp^{(N+1)}(m), XRp^{(N+1)}(m)]$; where $OPi[XLp^i(m), XRp^i(m)]$ indicates that OPi is determined as a function of input pair $XLp^i(m)$, $XRp^i(m)$. Correspondingly, the outputs of operation array 46 are $Xp^1(m)$, $Xp^2(m)$, $Xp^3(m)$, $Xp^4(m)$, ..., $Xp^{(N-2)}(m)$, $Xp^{(N-1)}(m)$, $Xp^N(m)$, and $Xp^{(N+1)}(m)$ (collectively designated $Xp^i(m)$).

For $i = 1$ to $i \leq N/2$, operations for each OPi of array 46 are determined in accordance with complex expression 1 (CE1) as follows:

$$Xp^i(m) = \frac{XLp^i(m) - XRp^i(m)}{\exp[-j2\pi(\tau_i+\ldots+\tau_{N/2})/f_m] - \exp[j2\pi(\tau_{((N/2)+1)}+\ldots+\tau_{(N-i+1)})/f_m]},$$

where exp[argument] represents a natural exponent to the power of the argument, and imaginary number j is the square root of -1. For $i > ((N/2)+1)$ to $i = N+1$, operations of operation array 46 are determined in accordance complex expression 2 (CE2) as follows:

$$Xp^i(m) = \frac{XLp^i(m) - XRp^i(m)}{\exp[j2\pi(\tau_{((N/2)+1)}+\ldots+\tau_{(i-1)})/f_m] - \exp[-j2\pi(\tau_{(N-i+2)}+\ldots+\tau_{N/2})/f_m]},$$

where exp[argument] represents a natural exponent to the power of the argument, and imaginary number j is the square root of -1. For $i = (N/2)+1$, neither CE1 nor CE2 is performed.

An example of the determination of the operations for $N = 4$ ($i=1$ to $i=N+1$) is as follows:

$i = 1$, CE1 applies as follows:

$$Xp^1(m) = \frac{XLp^1(m) - XRp^1(m)}{\exp[-j2\pi(\tau_1+\tau_2)/f_m] - \exp[j2\pi(\tau_3+\tau_4)/f_m]};$$

$i = 2 \leq (N/2)$, CE1 applies as follows:

$$Xp^2(m) = \frac{XLp^2(m) - XRp^2(m)}{\exp[-j2\pi(\tau_2)/f_m] - \exp[j2\pi(\tau_3)/f_m]};$$

$i = 3$: Not applicable, $(N/2) < i \leq ((N/2)+1)$;

$i = 4$, CE2 applies as follows:

$$Xp^4(m) = \frac{XLp^4(m) - XRp^4(m)}{\exp[j2\pi(\tau_3)/f_m] - \exp[-j2\pi(\tau_2)/f_m]};\text{ and,}$$

$i = 5$, CE2 applies as follows:

$$Xp^5(m) = \frac{XLp^5(m) - XRp^5(m)}{\exp[j2\pi(\tau_3+\tau_4)/f_m] - \exp[-j2\pi(\tau_1+\tau_2)/f_m]}.$$

Referring to FIGS. 1-3, each OPi of operation array 46 is defined to be representative of a different azimuthal position relative to reference axis R. The "center" operation, OPi where $i = ((N/2)+1)$, represents the location of the reference axis and source 12. For the example N=4, this center operation corresponds to $i = 3$.

This arrangement is analogous to the different interaural time differences associated with a natural binaural hearing system. In these natural systems, there is a relative position in each sound passageway within the ear that corresponds to a maximum "in phase" peak for a given sound source. Accordingly, each operation of array 46 represents a position corresponding to a potential azimuthal or angular position range for a sound source, with the center operation representing a source at the zero azimuth -- a source aligned with reference axis R. For an environment having a single source without noise or interference, determining the signal pair with the maximum strength may be sufficient to locate the source with little additional processing; however, in noisy or multiple source environments, further processing may be needed to properly estimate locations.

It should be understood that dual delay line 40 provides a two dimensional matrix of outputs with N+1 columns corresponding to $Xp^i(m)$, and M rows corresponding to each discrete frequency f_m of $Xp^i(m)$. This (N+1)xM matrix is determined for each short-term spectral analysis interval p. Furthermore, by subtracting $XRp^i(m)$ from $XLp^i(m)$, the denominator of each expression CE1, CE2 is arranged to provide a minimum value of $Xp^i(m)$ when the signal pair is "in-phase" at the given frequency f_m. Localization stage 70 uses this aspect of expressions CE1, CE2 to evaluate the location of source 14 relative to source 12.

Localization stage 70 accumulates P number of these matrices to determine the $Xp^i(m)$ representative of the position of source 14. For each column i, localization stage 70 performs a summation of the amplitude of $|Xp^i(m)|$ to the second power over frequencies f_m from m=1 to m=M. The summation is then multiplied by the inverse of M to find an average spectral energy as follows:

$$Xavgp^i = (1/M) \sum_{m=1}^{M} |Xp^i(m)|^2.$$

The resulting averages, $Xavgp^i$ are then time averaged over the P most recent spectral-analysis time frames indexed by p in accordance with:

$$X^i = \sum_{p=1}^{P} \gamma p \, Xavgp^i ,$$

where γp are empirically determined weighting factors. In one embodiment, the γp factors are preferably between 0.85^p and 0.90^p, where p is the short-term spectral analysis time frame index. The X^i are analyzed to determine the minimum value, $\min(X^i)$. The index i of $\min(X^i)$, designated "I," estimates the column representing the azimuthal location of source 14 relative to source 12.

It has been discovered that the spectral content of a desired signal from source 12, when approximately aligned with reference axis R1, can be estimated from $Xp^I(m)$. In other words, the spectral signal output by array 46 which most closely corresponds to the relative location of the "off-axis" source 14 contemporaneously provides a spectral representation of a signal emanating from source 12. As a result, the signal processing of dual delay line 40 not only facilitates localization of source 14, but also provides a spectral estimate of the desired signal with only minimal post-localization processing to produce a representative output.

Post-localization processing includes provision of a designation signal by localization stage 70 to conceptual "switch" 80 to select the output column $Xp^I(m)$ of the dual delay line 40. The $Xp^I(m)$ is routed by switch 80 to an inverse Discrete Fourier Transform algorithm (Inverse DFT) in stage 82 for conversion from a frequency domain signal representation to a discrete time domain signal representation denoted as s(k). The signal estimate s(k) is then converted by Digital to Analog (D/A) converter 84 to provide an output signal to output device 80.

Output device 80 amplifies the output signal from processor 30 with amplifier 92 and supplies the amplified signal to speaker 94 to provide the extracted signal from a source 12.

It has been found that interference from off-axis sources separated by as little as 2 degrees from the on axis source may be reduced or eliminated with the present invention -- even when the desired signal includes speech and the interference includes babble. Moreover, the present invention provides for the extraction of desired signals even when the interfering or noise signal is of equal or greater relative intensity. By moving sensors 22, 24 in tandem the signal selected to be extracted may correspondingly be changed. Moreover, the present invention may be employed in an environment having many sound sources in addition to sources 12, 14. In one alternative embodiment, the localization algorithm is configured to dynamically respond to relative positioning as well as relative strength, using automated learning techniques. In other embodiments, the present invention is adapted for use with highly directional microphones, more than two sensors to simultaneously extract multiple signals, and various adaptive amplification and filtering techniques known to those skilled in the art.

The present invention greatly improves computational efficiency compared to conventional systems by determining a spectral signal representative of the desired signal as part of the localization processing. As a result, an output signal characteristic of a desired signal from source 12 is determined as a function of the signal pair $XLp^I(m)$, $XRp^I(m)$ corresponding to the separation of source 14 from source 12. Also, the exponents in the denominator of CE1, CE2 correspond to phase difference of frequencies f_m resulting from the separation of source 12 from 14. Referring to the example of N=4 and assuming that I=1, this phase difference is $-2\pi(\tau_1+\tau_2)f_m$ (for delay line 42) and $2\pi(\tau_3+\tau_4)f_m$ (for delay line 44) and corresponds to the separation of the representative location of off-axis source 14 from the on-axis source 12 at i=3. Likewise the time increments, $\tau_1+\tau_2$ and $\tau_3+\tau_4$,

correspond to the separation of source 14 from source 12 for this example. Thus, processor 30 implements dual delay line 40 and corresponding operational relationships CE1, CE2 to provide a means for generating a desired signal by locating the position of an interfering signal source relative to the source of the desired signal.

It is preferred that τ_i be selected to provide generally equal azimuthal positions relative to reference axis R. In one embodiment, this arrangement corresponds to the values of τ_i changing about 20% from the smallest to the largest value. In other embodiments, τ_i are all generally equal to one another, simplifying the operations of array 46. Notably, the pair of time increments in the numerator of CE1, CE2 corresponding to the separation of the sources 12 and 14 become approximately equal when all values τ_i are generally the same.

Processor 30 may be comprised of one or more components or pieces of equipment. The processor may include digital circuits, analog circuits, or a combination of these circuit types. Processor 40 may be programmable, an integrated state machine, or utilize a combination of these techniques. Preferably, processor 40 is a solid state integrated digital signal processor circuit customized to perform the process of the present invention with a minimum of external components and connections. Similarly, the extraction process of the present invention may be performed on variously arranged processing equipment configured to provide the corresponding functionality with one or more hardware modules, firmware modules, software modules, or a combination thereof. Moreover, as used herein, "signal" includes, but is not limited to, software, firmware, hardware, programming variable, communication channel, and memory location representations.

Referring to FIG. 4A, one application of the present invention is depicted as hearing aid system 110. System 110 includes eyeglasses G with microphones 122 and 124 fixed to glasses G and displaced from one another. Microphones 122, 124 are operatively coupled to hearing aid processor 130. Processor 130 is operatively coupled

to output device 190. Output device 190 is positioned in ear E to provide an audio signal to the wearer.

Microphones 122, 124 are utilized in a manner similar to sensors 22, 24 of the embodiment depicted by FIGS 1-3. Similarly, processor 130 is configured with the signal extraction process depicted in of FIGS. 1-3. Processor 130 provides the extracted signal to output device 190 to provide an audio output to the wearer. The wearer of system 110 may position glasses G to align with a desired sound source, such as a speech signal, to reduce interference from a nearby noise source off axis from the midpoint between microphones 122, 124. Moreover, the wearer may select a different signal by realigning with another desired sound source to reduce interference from a noisy environment.

Processor 130 and output device 190 may be separate units (as depicted) or included in a common unit worn in the ear. The coupling between processor 130 and output device 190 may be an electrical cable or a wireless transmission. In one alternative embodiment, sensors 122, 124 and processor 130 are remotely located and are configured to broadcast to one or more output devices 190 situated in the ear E via a radio frequency transmission or other conventional telecommunication method.

FIG. 4B shows a voice recognition system 210 employing the present invention as a front end speech enhancement device. System 210 includes personal computer C with two microphones 222, 224 spaced apart from each other in a predetermined relationship. Microphones 222, 224 are operatively coupled to a processor 230 within computer C. Processor 230 provides an output signal for internal use or responsive reply via speakers 294a, 294b or visual display 296. An operator aligns in a predetermined relationship with microphones 222, 224 of computer C to deliver voice commands. Computer C is configured to receive these voice commands, extracting the desired voice command from a noisy environment in accordance with the process system of FIGS. 1-3.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

EXPERIMENTAL SECTION

The following experimental results are provided as nonlimiting examples, and should not be construed to restrict the scope of the present invention.

A Sun Sparc-20 workstation was programmed to emulate the signal extraction process of the present invention. One loudspeaker (L1) was used to emit a speech signal and another loudspeaker (L2) was used to emit babble noise in a semi-anechoic room. Two microphones of a conventional type where positioned in the room and operatively coupled to the workstation. The microphones had an inter-microphone distance of about 15 centimeters and were positioned about 3 feet from L1. L1 was aligned with the midpoint between the microphones to define a zero degree azimuth. L2 was placed at different azimuths relative to L1 approximately equidistant to the midpoint between L1 and L2.

Referring to FIG. 5, a clean speech of a sentence about two seconds long is depicted, emanating from L1 without interference from L2. FIG. 6 depicts a composite signal from L1 and L2. The composite signal includes babble noise from L2 combined with the speech signal depicted in FIG. 5. The babble noise and speech signal are of generally equal intensity (0dB) with L2 placed at a 60 degree azimuth relative to L1. FIG. 7 depicts the signal recovered from the composite signal of FIG. 6. This signal is nearly the same as the signal of FIG. 5.

FIG. 8 depicts another composite signal where the babble noise is 30dB more intense than the desired signal of FIG. 5. Furthermore, L2 is placed at only a 2 degree azimuth relative to L1. FIG. 9 depicts the signal recovered from the composite signal of FIG. 8, providing a clearly intelligible representation of the signal of FIG. 5 despite the greater intensity of the babble noise from L2 and the nearby location.

While the invention has been illustrated and described in detail in the drawings and foregoing description, the same is to be considered as illustrative and not restrictive in character, it being understood that only the preferred embodiment has been shown and described and that all changes and modifications that come within the spirit of the invention are desired to be protected.

We claim:

1. A method of signal processing, comprising:

(a) detecting an acoustic excitation at a first location to provide a corresponding first signal and at a second location to provide a corresponding second signal, the excitation including a desired acoustic signal from a first source and an interfering acoustic signal from a second source spaced apart from the first source;

(b) localizing the second source relative to the first source as a function of the first and second signals; and

(c) generating a characteristic signal representative of the desired acoustic signal during performance of said localizing.

2. The method of claim 1, wherein the characteristic signal corresponds to spectral content of the desired acoustic signal and further comprising providing an output signal representative of the desired acoustic signal as a function of the characteristic signal.

3. The method of claim 1, wherein said localizing includes:

(b1) delaying each of the first and second signals by a number of time intervals to provide a number of delayed first signals and a number of delayed second signals; and

(b2) determining a time interval representative of separation of the first source from the second source, the characteristic signal being a function of the time interval.

4. The method of claim 1, wherein said localizing includes:

(b1) delaying each of the first and second signals by a number of time intervals to provide a number of delayed first signals and a number of delayed second signals; and

(b2) establishing a signal pair, the signal pair having a first member from the delayed first signals and a second member from the delayed second members, the characteristic signal being determined from the signal pair.

5. The method of claim 1, further comprising providing an output signal representative of the desired acoustic signal, and wherein the desired acoustic signal includes speech and the output signal is provided by a hearing aid device.

6. The method of claim 1, wherein said localizing further includes:

(b1) converting the first and second signals from an analog representation to a discrete representation;

(b2) transforming the first and second signals from a time domain representation to a frequency domain representation;

(b3) delaying each of the first and second signals by a number of time intervals to provide a number of delayed first signals and a number of delayed second signals; and

(b4) establishing a first time increment and a signal pair each representative of separation of the first source from the second source, the signal pair having a first member from the delayed first signals and a second member from the delayed second members.

7. The method of claim 6, wherein the characteristic signal corresponds to a fraction with a numerator determined from at least the first and second members, and a denominator determined from at least the first time increment.

8. The method of claim 6, wherein said generating further includes:

(c1) determining the characteristic signal from the signal pair and the first time increment, the characteristic signal being representative of spectral content of the desired acoustic signal;

(c2) transforming the characteristic signal from a frequency domain representation to a time domain representation;

(c3) converting the characteristic signal from a discrete representation to an analog representation; and

(c4) providing an audio output signal representative of the desired acoustic signal as a function of the characteristic signal.

9. The method of claim 8, further comprising establishing a second time increment corresponding to separation of the first source from the second source by comparing the delayed first and second signals, and

wherein the first time increment corresponds to a first phase difference, the second time increment corresponds to a second phase difference, and the characteristic signal includes a spectral representation determined from at least the first and second phase differences.

10. The method of claim 1, wherein the desired acoustic signal has an intensity greater than the interfering acoustic signal when the first and second sources are each generally equidistant from a midpoint between the first and second locations.

11. The method of claim 1, wherein separation of the second source is within five degrees of the first source relative to a zero degree azimuthal reference axis intersecting the first source and a midpoint situated between the first and second locations.

12. The method of claim 1, further comprising:

(d) establishing a number of location signals, each corresponding to a different location relative to the first source; and

(e) selecting the characteristic signal from the location signals, the characteristic signal being representative of location of the second source relative to the first source, the characteristic signal including a spectral representation of the desired acoustic signal.

13. A signal processing system, comprising:

(a) a first sensor at a first location configured to provide a first signal corresponding to an acoustic signal, said acoustic signal including a desired signal emanating from a selected source and noise emanating from a noise source;

(b) a second sensor at a second location configured to provide a second signal corresponding to said acoustic signal;

(c) a signal processor responsive to said first and second signals to generate a discrete first spectral signal corresponding to said first signal and a discrete second spectral signal corresponding to said second signal, said processor being configured to delay said first and second spectral signals by a number of time intervals to generate a number of delayed first signals and a number of delayed second signals and provide a time increment signal, said time increment signal corresponding to separation of the selected source from the noise source, and said processor being further configured to generate an output signal as a function of said time increment signal; and

(d) an output device responsive to said output signal to provide an output representative of said desired signal.

14. The system of claim 13, wherein said first and second sensors each include a microphone and said output device includes an audio speaker.

15. The system of claim 13, wherein said processor includes an analog to digital conversion circuit configured to provide said discrete first spectral signal.

16. The system of claim 13, wherein generation of said first and second spectral signals includes execution of a discrete fourier transform algorithm.

17. The system of claim 13, wherein said first and second sensors are configured for movement to select said desired signal in accordance with position of said first and second sensors, said first and second sensors being configured to be spatially fixed relative to each other.

18. The system of claim 13, wherein each of said delayed first signals correspond to one of a number of first taps from a first delay line, and each of said delayed second signals correspond to one of a number of second taps from a second delay line.

19. The system of claim 18, wherein determination of said output signal corresponds to:

said first and second delay lines being configured in a dual delay line configuration;

said discrete first spectral signal being input to said first delay line and said discrete second spectral signal being input to said second delay line; and

each of said first taps, said second taps, and said first and second spectral signals being arranged as a number of signal pairs, said signal pairs including a first portion of signal pairs and a second portion of signal pairs, said processor being configured to perform a first operation on each of said signal pairs of said first portion as a function of said time intervals, said processor being configured to perform a second operation on each of said signal pairs of said second portion as a function of said time intervals, said first operation being different from said second operation.

20. A signal processing system, comprising:

(a) a first sensor configured to provide a first signal corresponding to an acoustic excitation, said excitation including a first acoustic signal from a first source and a second acoustic signal from a second source displaced from the first source;

(b) a second sensor displaced from said first sensor and configured to provide a second signal corresponding to said excitation;

(c) a processor responsive to said first and second sensor signals, said processor including a means for generating a desired signal having a spectrum representative of said first acoustic signal; and

(d) an output means for generating a sensory output in response to said desired signal.

21. The system of claim 20, wherein said first and second sensors each include a microphone and said output device includes an audio speaker.

22. The system of claim 20, wherein said generating means includes executing a discrete fourier transform algorithm.

23. The system of claim 20, wherein said processor includes an analog to digital conversion circuit and a digital to analog conversion circuit.

24. The system of claim 20, wherein each of said delayed first signals correspond to one of a number of first taps from a first delay line, and each of said delayed second signals correspond to one of a number of second taps from a second delay line.

25. The system of claim 20, wherein said first and second sensors are configured for movement to select said desired signal in accordance with position of said first and second

sensors, said first and second sensors being configured to be spatially fixed relative to each other.

26. A method of signal processing, comprising:

(a) positioning a first and second sensor relative to a first signal source, the first and second sensor being spaced apart from each other, and a second signal source being spaced apart from the first signal source;

(b) providing a first signal from the first sensor and a second signal from the second signal, the first and second signals each being representative of a composite acoustic signal including a desired signal from the first signal source and an unwanted signal from the second signal source;

(c) establishing a number of spectral signals from the first and second signals as a function of a number of frequencies, each of the spectral signals representing a different position relative to the first signal source;

(d) determining a member of the spectral signals representative of position of the second signal source; and

(e) generating an output signal from the member, the output signal being representative of spectral content of the first signal.

27. The method of claim 26, wherein the member is determined as a function of a phase difference value for a number of frequencies delayed by a first amount and a second amount.

28. The method of claim 26, wherein the desired signal includes speech and the output signal is provided by a hearing aid device.

29. The method of claim 26, further comprising repositioning the first and second sensors to extract a third signal from a third signal source.

30. The method of claim 26, wherein said establishing includes:

(a1) delaying each of the first and second signals by a number of time intervals to generate a number of delayed first signals and a number of delayed second signals; and

(a2) comparing each of the delayed first signals to a corresponding one of the delayed second signals, each of the spectral signals being a function of at least one of the delayed first and second signals.

ABSTRACT OF THE DISCLOSURE

A desired acoustic signal is extracted from a noisy environment by generating a signal representative of the desired signal with a processor for a hearing aid device. The processor receives binaural signals from two microphones at different locations. The binaural inputs to the processor are converted from analog to digital format and then submitted to a discrete Fourier transform process to generate discrete spectral signal representations. The spectral signals are delayed by a number of time intervals in a dual delay line to provide a number of intermediate signals, each corresponding to a different position relative to a desired signal source. Location of the noise source is determined and the spectral content of the desired signal is determined from the intermediate signal corresponding to the noise source location. Inverse transformation of the selected intermediate signal followed by digital to analog conversion provides an output signal representative of the desired signal.

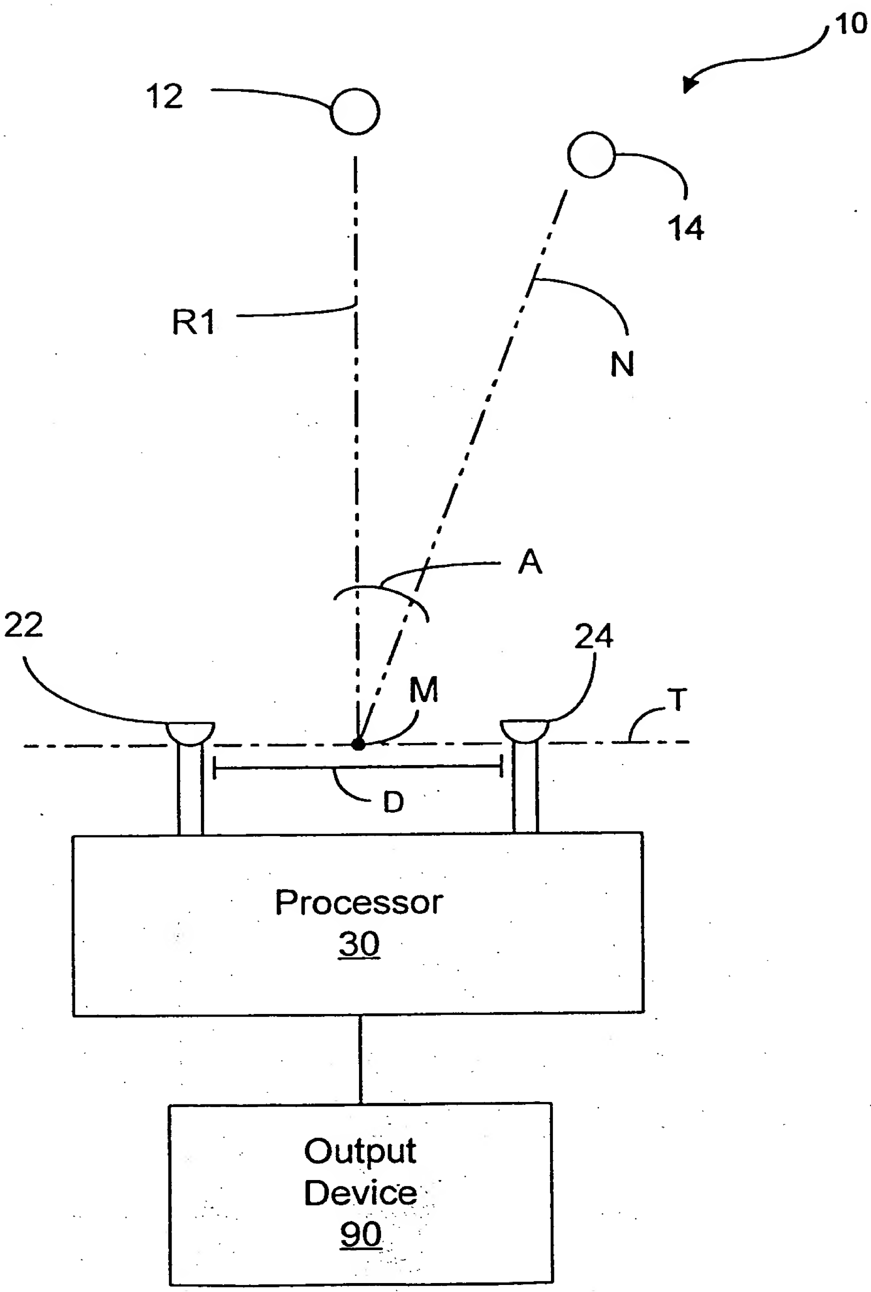
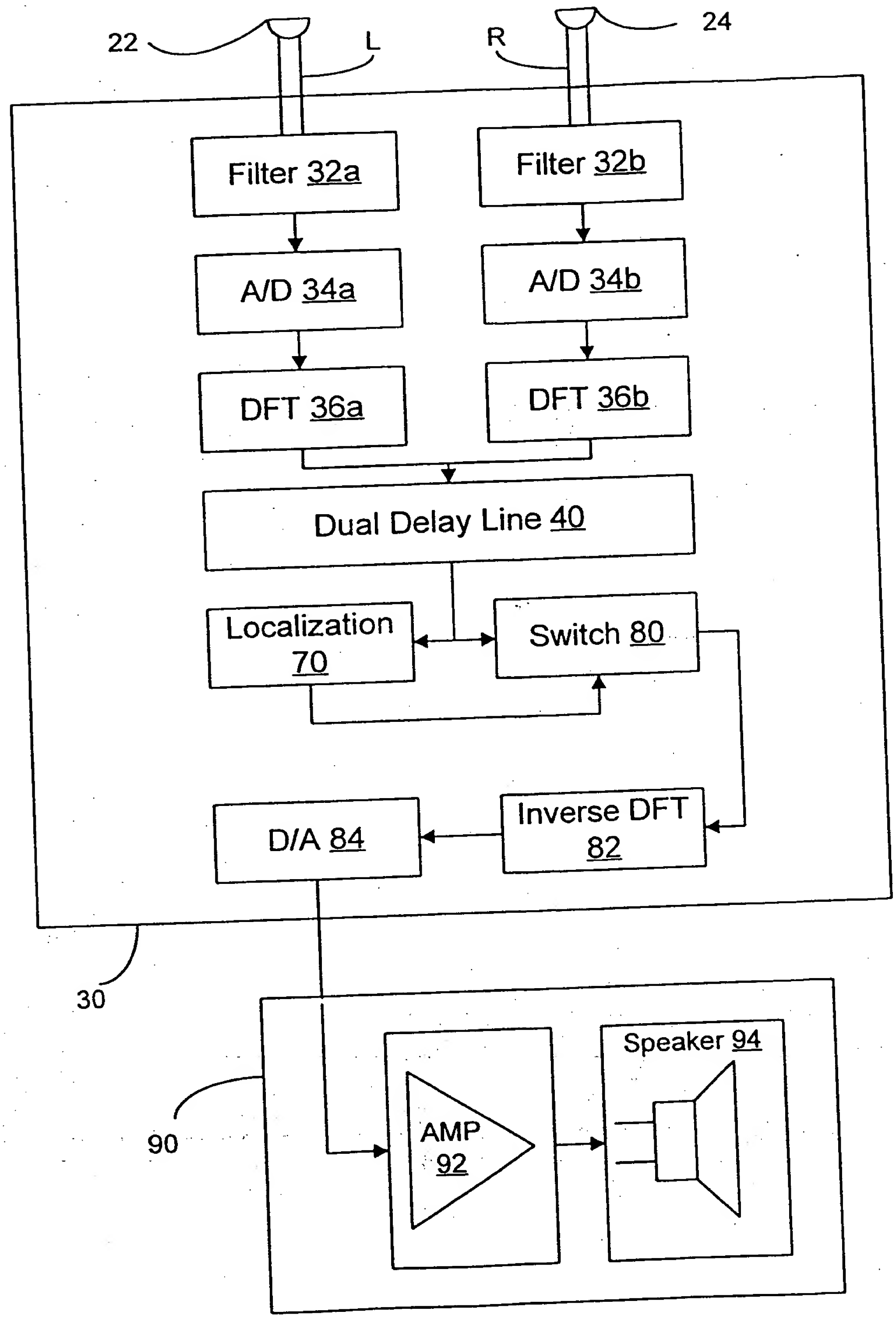
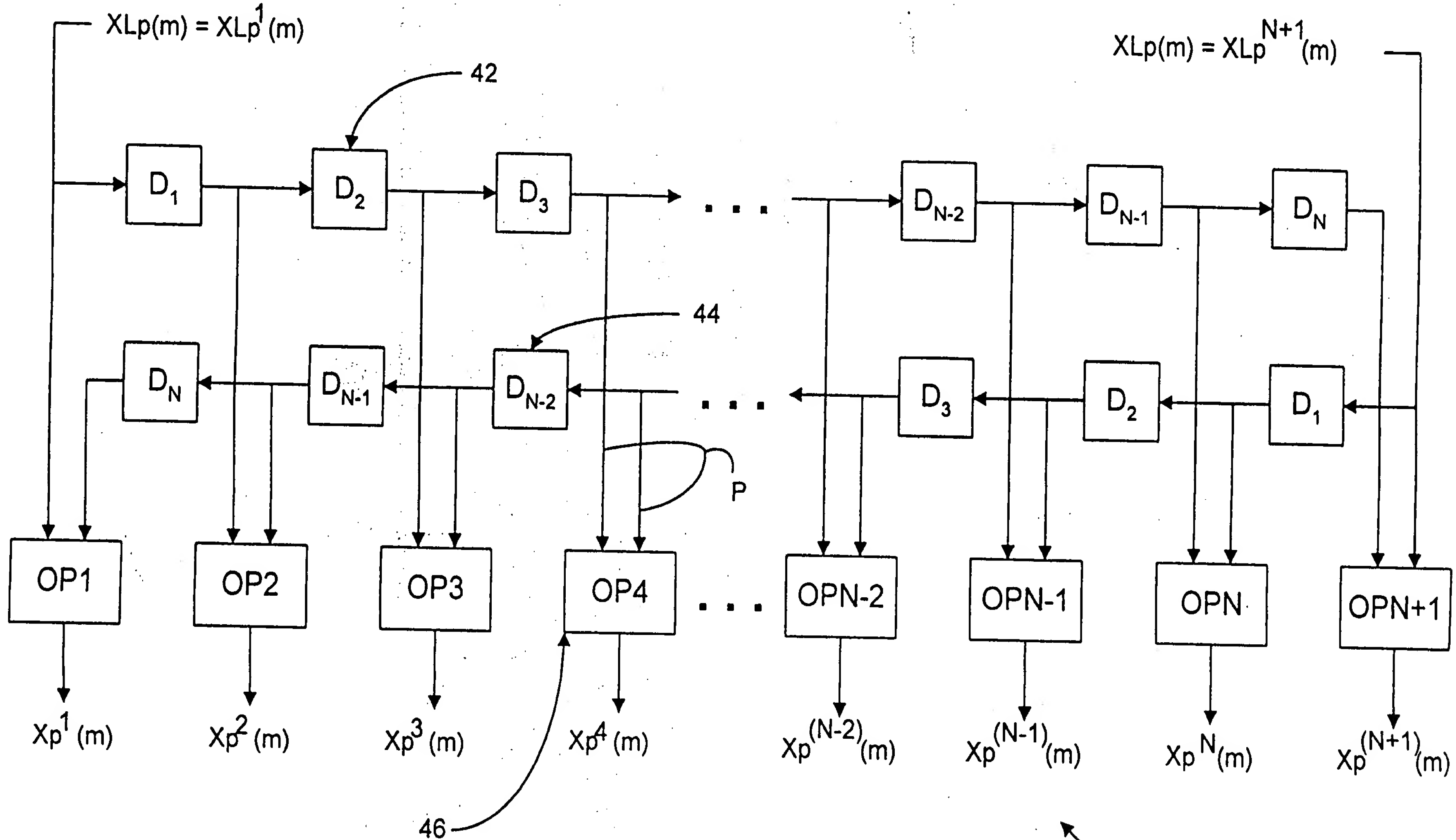


FIG. 1



22
L
R
24
Filter 32a
Filter 32b
A/D 34a
A/D 34b
DFT 36a
DFT 36b
Dual Delay Line 40
Localization 70
Switch 80
Inverse DFT 82
D/A 84
30
90
AMP 92
Speaker 94

FIG. 2



XLp(m) = XLp1(m)
XLp(m) = XLpN+1(m)
42
44
46
40
P
D1
D2
D3
DN-2
DN-1
DN
D1
D2
D3
DN-2
DN-1
DN
OP1
OP2
OP3
OP4
OPN-2
OPN-1
OPN
OPN+1
Xp1 (m)
Xp2 (m)
Xp3 (m)
Xp4 (m)
Xp(N-2)(m)
Xp(N-1)(m)
Xp N(m)
Xp(N+1)(m)

FIG. 3

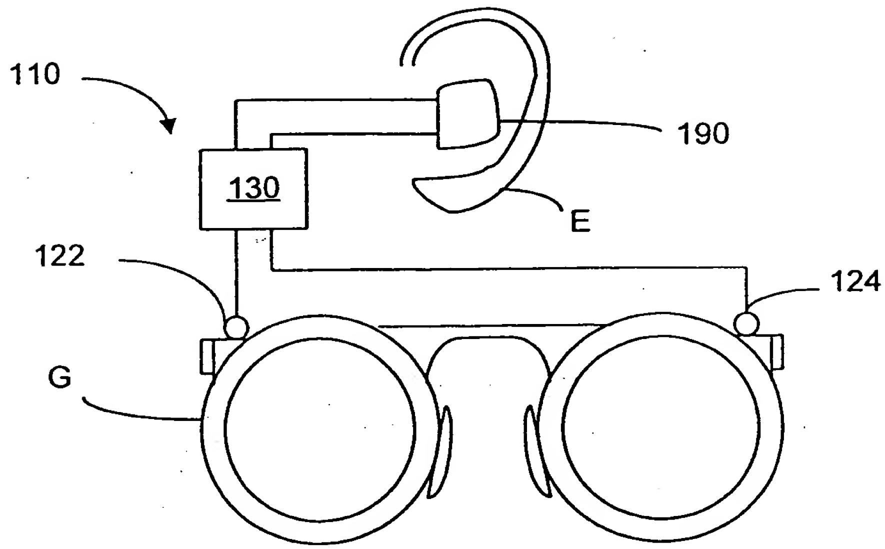


FIG. 4A

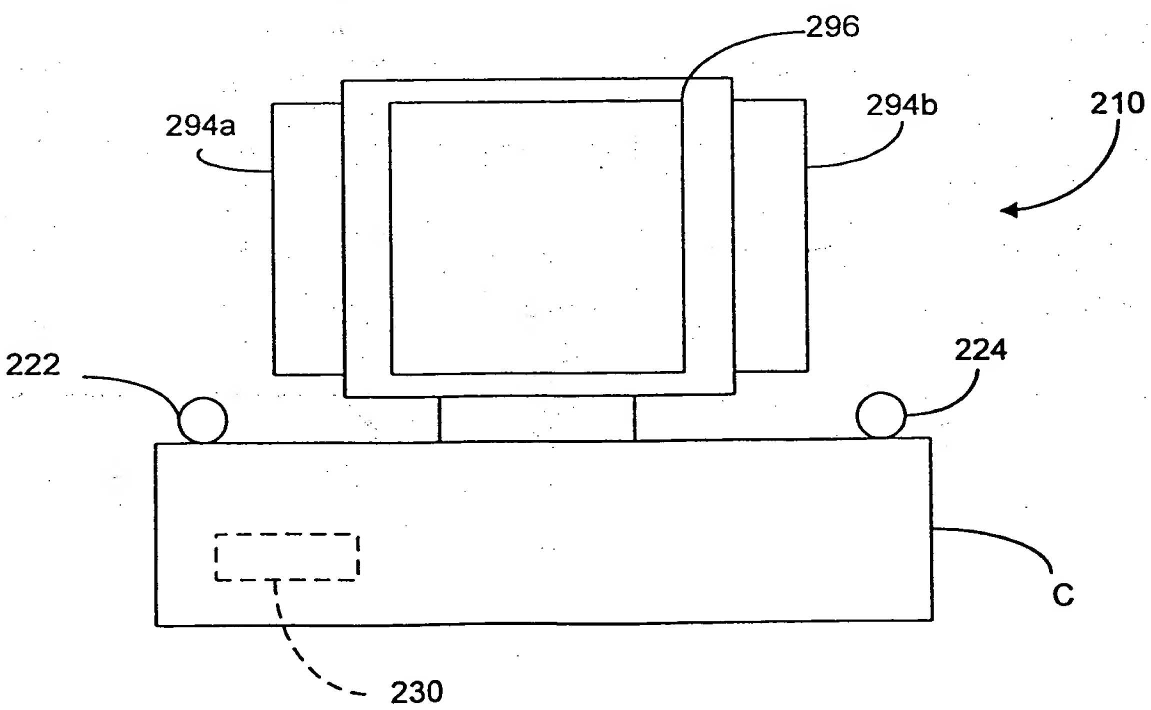
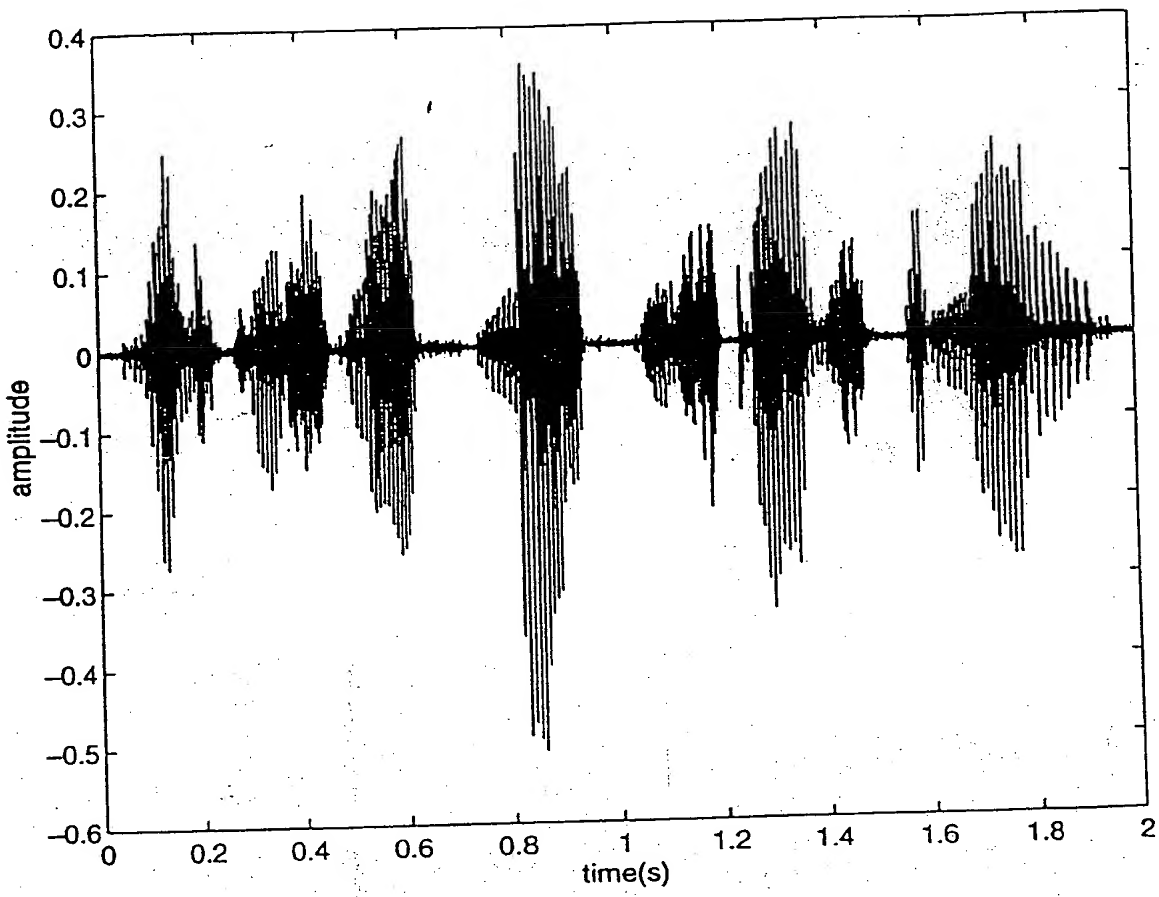
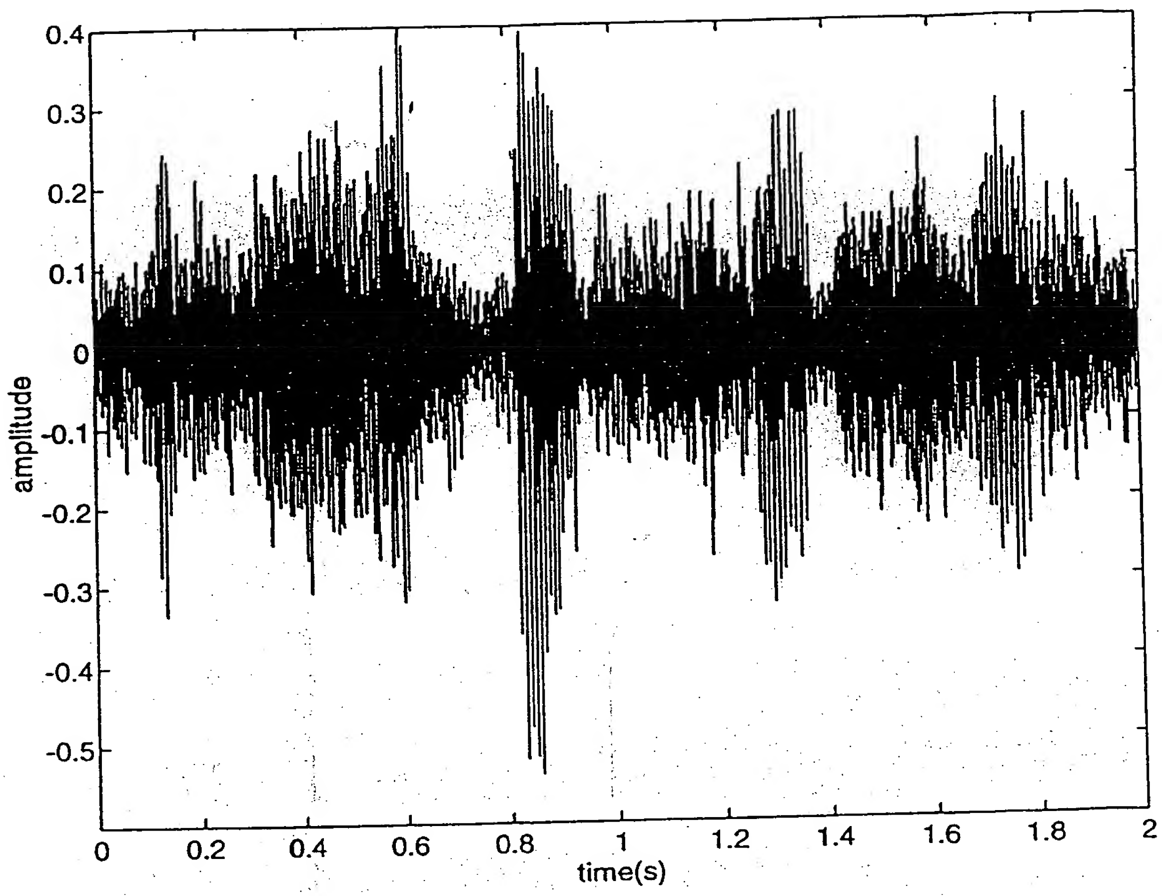


FIG. 4B



0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

FIG. 5



0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

FIG. 6

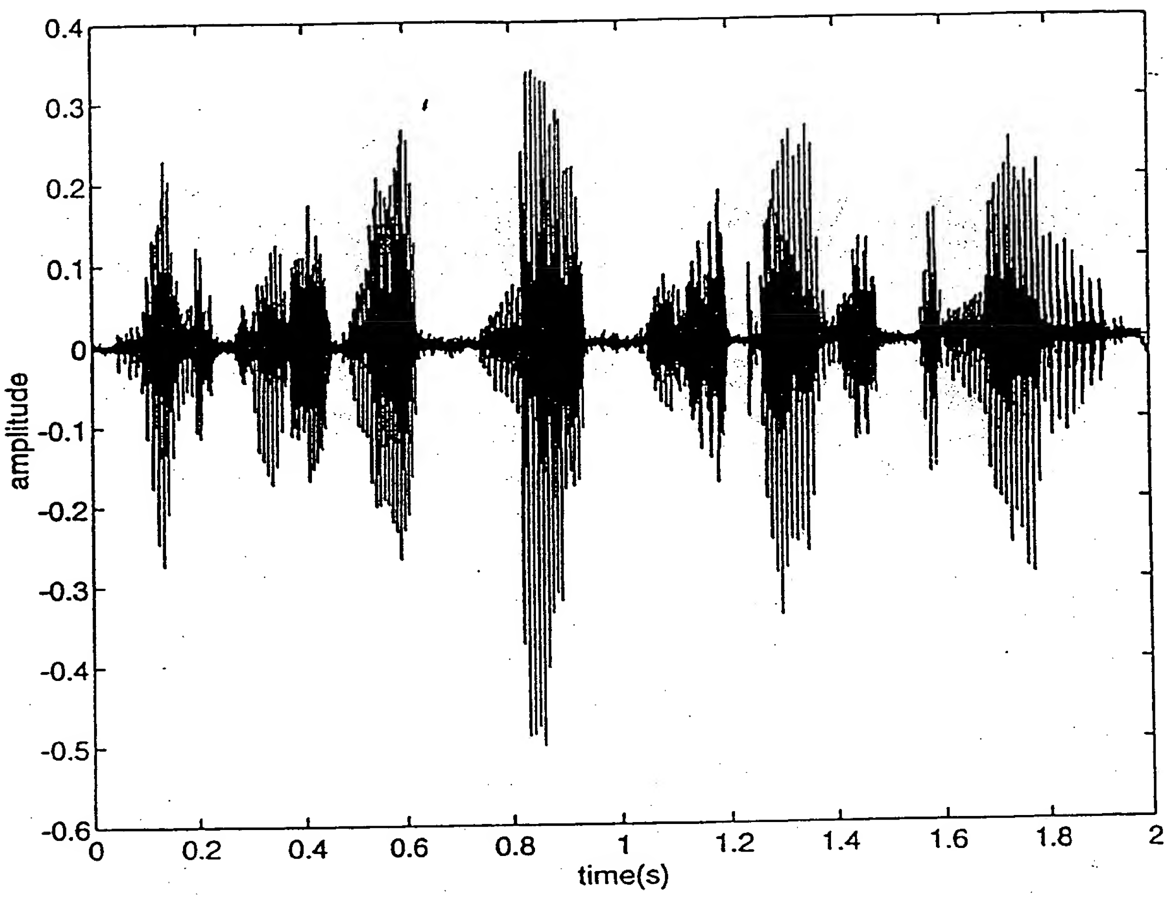

0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

FIG. 7

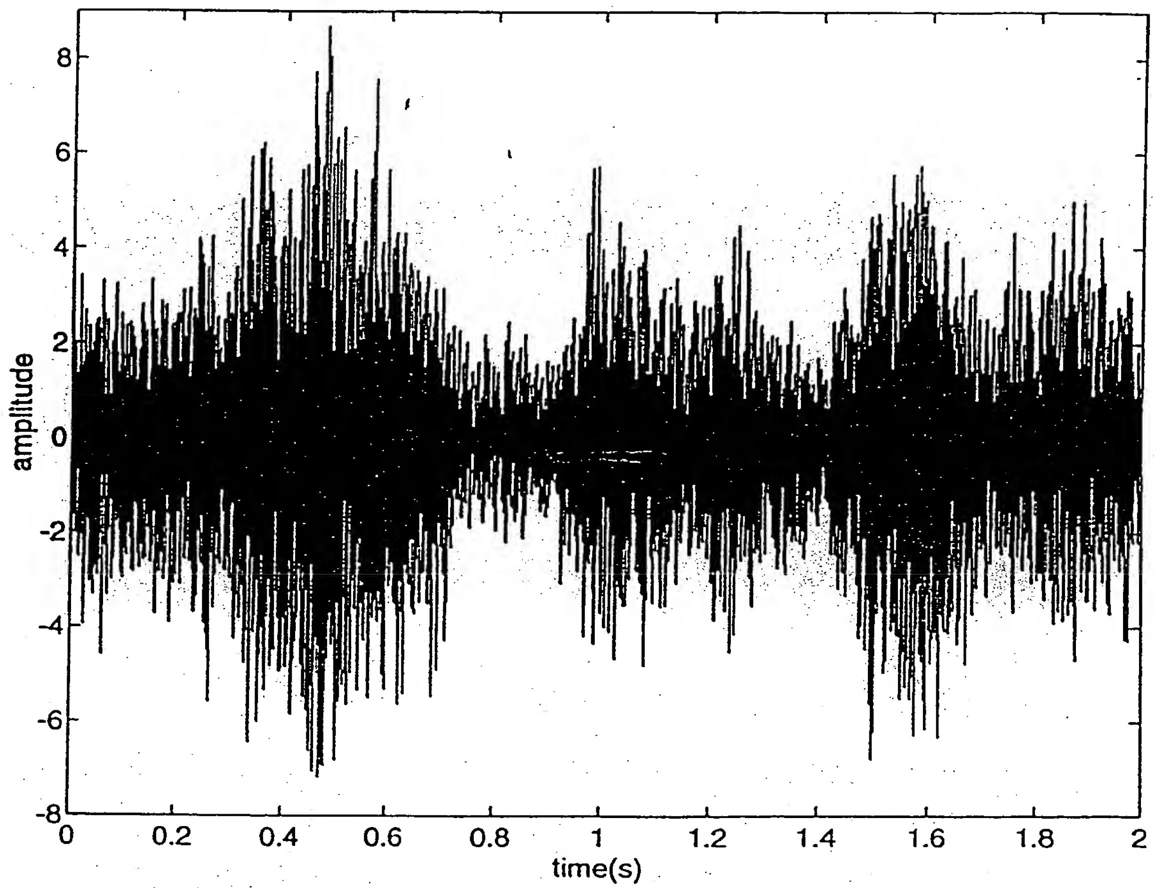
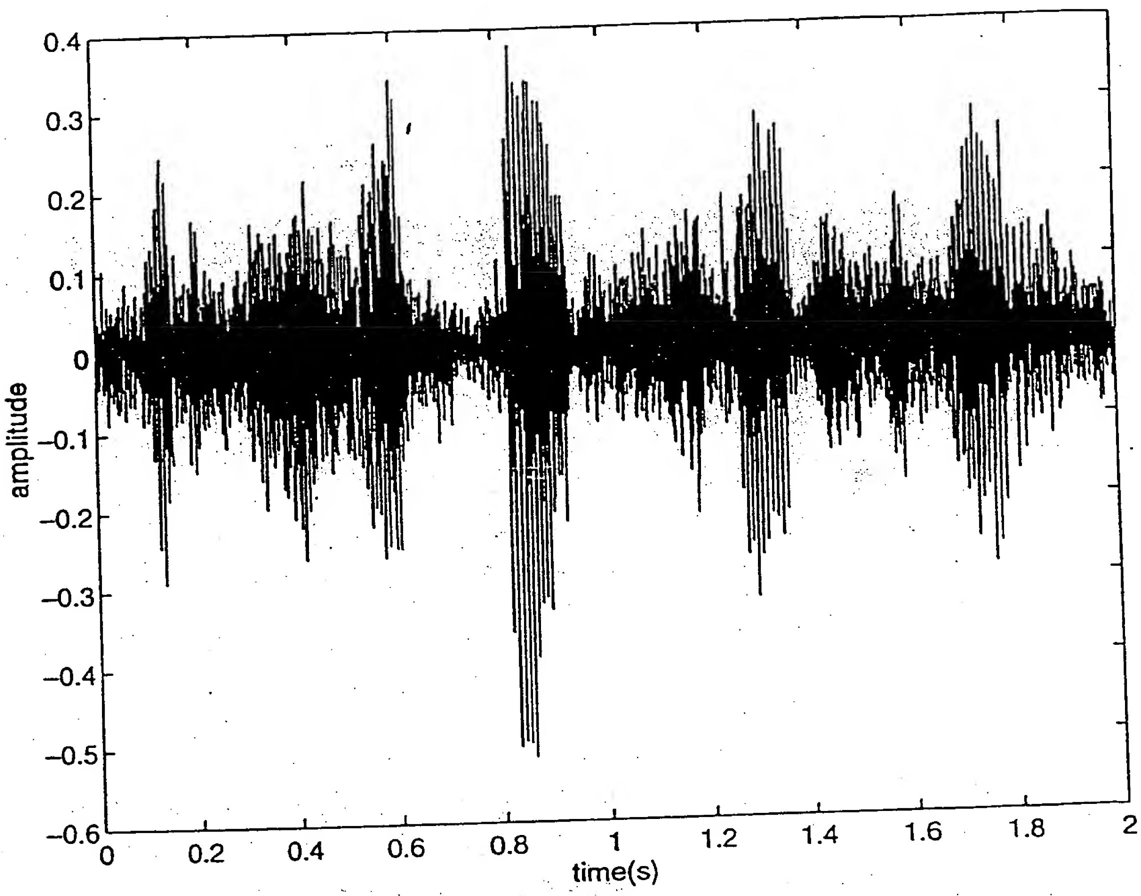


FIG. 8



0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

FIG. 9

Patent Claims

1.-33. (Canceled).

34. (Previously presented) A method of signal processing, comprising:

(a) detecting an acoustic excitation at both a first location to provide a corresponding first signal and at a second location to provide a corresponding second signal, the excitation being a composite of a desired acoustic signal from a first source and an interfering acoustic signal from a second source spaced apart from the first source;

(b) determining location of the second source relative to the first source as a function of the first and second signals, which includes delaying each of the first and second signals by several time intervals to provide several delayed first signals and several delayed second signals and providing a time increment representative of separation of the first source from the second source; and

(c) generating a characteristic signal representative of the desired acoustic signal during performance of said determining, the characteristic signal being a function of the time increment.

35. (Previously presented) The method of claim 34, wherein the characteristic signal corresponds to spectral content of the desired acoustic signal and further comprising providing an output signal representative of the desired acoustic signal as a function of the characteristic signal.

36. (Canceled).

37. (Currently amended) The method of claim 34, wherein said determining includes:

establishing a signal pair, the signal pair having a first member from the delayed first signals and a second member from the delayed second signals, the characteristic signal being determined from the signal pair.

38. (Previously presented) The method of claim 34, further comprising providing an output signal representative of the desired acoustic signal, and wherein the desired acoustic signal includes speech and the output signal is provided by a hearing aid device.

39. (Currently amended) The method of claim 34, wherein said determining further includes:

(b1) converting the first and second signals from an analog representation to a discrete representation;

(b2) transforming the first and second signals from a time domain representation to a frequency domain representation; <u>and</u>

(b3) establishing a signal pair representative of separation of the first source from the second source, the signal pair having a first member from the delayed first signals and a second member from the delayed second signals.

40. (Currently amended) The method of claim 39, wherein the characteristic signal corresponds to a fraction with a numerator determined from at least the first and second members, and a denominator determined from at least the time increment.

41. (Previously presented) The method of claim 39, wherein said generating further includes:

(c1) determining the characteristic signal from the signal pair and the first time increment, the characteristic signal being representative of spectral content of the desired acoustic signal;

(c2) transforming the characteristic signal from a frequency domain representation to a time domain representation; and

(c3) providing an audio output signal representative of the desired acoustic signal as a function of the characteristic signal.

42. (Currently amended) The method of claim 41, further comprising establishing a further time increment corresponding to separation of the first source from the second source by comparing the delayed first and second signals, and

wherein the time increment corresponds to a first phase difference, the further time increment corresponds to a second phase difference, and the characteristic signal includes a spectral representation determined from at least the first and second phase differences.

43. (Canceled).

44. (Previously presented) The method of claim 34, wherein separation of the second source is within five degrees of the first source relative to a zero degree azimuthal

reference axis intersecting the first source and a midpoint situated between the first and second locations.

45. (Previously presented) The method of claim 34, further comprising;

(d) establishing a number of location signals each corresponding to a different location relative to the first source; and

(e) selecting the characteristic signal from the location signals, the characteristic signal being representative of the location of the second source relative to the first source, the characteristic signal including a spectral representation of the desired acoustic signal.

46. (Previously presented) A method of signal processing, comprising:

(a) detecting an acoustic excitation at a first location to provide a corresponding first signal and at a second location to provide a corresponding second signal, the excitation being a composite of a desired acoustic signal from a first source and an interfering acoustic signal from a second source spaced apart from the first source;

(b) localizing the second source relative to the first source as a function of the first and second signals, said localizing including establishing a number of location signals each corresponding to a different location relative to the first source, delaying each of the first and second signals by a number of time intervals to provide a number of delayed first signals and a number of delayed second signals, and establishing a signal pair that has a first member from the delayed first signals and a second member from the delayed second signals; and

(c) generating a characteristic signal from the location signals, wherein the characteristic signal includes a spectral representation of the desired acoustic signal from the first source, corresponds to position of the second source, and is determined from the signal pair.

47. (Previously presented) The method of claim 46, further comprising providing an output signal representative of the desired acoustic signal as a function of the characteristic signal.

48. (Currently amended) The method of claim 46, wherein said localizing includes:

determining a time increment representative of separation of the first source from the second source, the characteristic signal being a function of the time increment.

49. (Canceled).

50. (Previously presented) The method of claim 46, further comprising providing an output signal representative of the desired acoustic signal, and wherein the desired acoustic signal includes speech and the output signal is provided by a hearing aid device.

51. (Currently amended) The method of claim 46, wherein said localizing further includes:

(b1) converting the first and second signals from an analog representation to a discrete representation;

(b2) transforming the first and second signals from a time domain representation to a frequency domain representation; and

(b3) establishing a first time increment and a signal pair each representative of separation of the first source from the second source, the signal pair having a first member from the delayed first signals and a second member from the delayed second signals.

52. (Previously presented) The method of claim 51, wherein the characteristic signal corresponds to a fraction with a numerator determined from at least the first and second members, and a denominator determined from at least the first time increment.

53. (Previously presented) The method of claim 51, wherein said generating further includes:

(c1) determining the characteristic signal from the signal pair and the first time increment;

(c2) transforming the characteristic signal from a frequency domain representation to a time domain representation; and

(c3) providing an audio output signal representative of the desired acoustic signal as a function of the characteristic signal.

54. (Previously presented) The method of claim 53, further comprising establishing a second time increment corresponding to separation of the first source from the second source by comparing the delayed first signals and delayed second signals, and

wherein the first time increment corresponds to a first phase difference, the second time increment corresponds to a second phase difference, and the spectral representation of the characteristic signal is determined from at least the first and second phase differences.

55. (Canceled).

56. (Previously presented) The method of claim 1, wherein separation of the second source is within five degrees of the first source relative to a zero degree azimuthal reference axis intersecting the first source and a midpoint situated between the first and second locations.

57. (New) The method of claim 34, wherein the characteristic signal corresponds to a fraction with a numerator determined from a difference between a first member of the delayed first signals and a second member of the delayed second signals, and a denominator determined from at least the time increment.

58. (New) The method of claim 57, which includes providing the delayed first signals from a first multistage delay line and the delayed second signals from a second multistage delay line, the first member being output by a stage of the first delay line corresponding to the location of the second source and the second member being output by a stage of the second delay line corresponding to the location of the second source, and a different stage

of each of the first delay line and the second delay line corresponding to location of the first source.

59. (New) The method of claim 58, wherein the difference is representative of a minimized interfering acoustic signal level and provides the characteristic signal representative of spectral content of the desired acoustic signal.

60. (New) The method of claim 46, wherein the generating includes determining the characteristic signal as a fraction with a numerator being a function of a difference between one of the delayed first signals and one of the delayed second signals, the difference being representative of a minimized interfering acoustic signal level, and the fraction having a denominator determined as a function of at least the first time increment.

61. (New) A method of signal processing, comprising:

detecting an acoustic excitation at both a first location to provide a corresponding first signal and at a second location to provide a corresponding second signal, the excitation being a composite of a desired acoustic signal from a first source and an interfering acoustic signal from a second source spaced apart from the first source;

incrementally delaying the first signal to provide a number of delayed first signals and the second signal to provide a number of delayed second signals, a number of different pairings of the delayed first signals and the delayed second signals representing different locations;

localizing the second source relative to one of the different locations as a function of a difference between the members of a corresponding one of the different pairings; and

generating a characteristic signal representative of spectral content of the desired acoustic signal based on the difference and a time increment corresponding to distance separating the first source and the second source.

62. (New) A method of signal processing, comprising:

detecting an acoustic excitation at both a first location to provide a corresponding first signal and at a second location to provide a corresponding second signal, the excitation being a composite of a desired acoustic signal from a first source and an interfering acoustic signal from a second source spaced apart from the first source;

selecting the desired acoustic signal by positioning a reference axis relative to the first source;

localizing the second source relative to the reference axis as a function of the first and second signals; and

generating a characteristic signal representative of the desired acoustic signal during performance of said localizing.

63. (New) The method of claim 62, which includes:

defining the reference axis relative to the first location and the second location; and

moving the reference axis to select a different acoustic signal.

64. (New) The method of claim 63, wherein the detecting the acoustic excitation is performed with a first sensor at the first location and a second sensor at the second location.

65. (New) The method of claim 63, wherein the method is performed with a hearing aid.

66. (New) The method of claim 63, wherein:

the localizing includes establishing a number of delayed first signals each corresponding to a different one of a number of first delay stages of a first delay line and a number of delayed second signals each corresponding to a different one of a number of second delay stages of a second delay line; and

the generating includes determining the characteristic signal as a function of a fraction with a numerator corresponding to a difference between one output of the first delay stages and one output of the second delay stages and a denominator corresponding to a time increment representative of a distance separating the first source and the second source.

University of Illinois
at Urbana-Champaign

Office of the
Vice Chancellor for Research
and Graduate College

Research and Technology Management Office
Fourth Floor Swanlund Building
601 East John Street
Champaign, IL 61820

217 333-7862
217 244-3716 *fax*
email: delkranz@uiuc.edu

May 9, 1996

RECEIVED

MAY 10 1996

WOODARD, EMHARDT, NAUGHTON
MORIARTY & McNETT

Kenneth Gandy, Esquire
Woodard, Emhardt, Naughton, Moriarty & McNett
Bank One Center Tower,
111 Monument Circle Suite 3700
Indianapolis, IN 46204

Re: New Invention Disclosure titled "New Dual-Microphone-Based Signal Extraction Algorithm That Can Acquire An Acoustic Signal Faithfully In The Presence Of An Intense Noise Originating From a Nearby Source", Al Feng et al., T96057

Dear Ken:

As we discussed the other day, enclosed is a copy of the invention disclosure and background information for the "hearing aid" invention.

As soon as you have had a chance to review the enclosed invention, please give Mel a call with your estimate for the work to be done, and for some additional background information. As soon as we have the estimate, we'll initiate the Letter of Referral for our legal counsel to authorize the work.

Let me know if you require anything further.

Sincerely yours,

Delphine Kranz

Delphine M. Kranz
Assistant Vice Chancellor for Research

DMK/ms
Enc.
c: A. Feng
M. DeGeeter
J. Quirk
J. Jonas

RECEIVED
MAY 09 1997
VICE CHAN FOR RES

CONFIDENTIAL

INVENTION DISCLOSURE

An invention may be any new product, process (method or use) or composition, or improvement thereof. To be patentable or protectable an invention must be novel (as compared to the existing state of the art), useful and non-obvious to a person with ordinary skill in the art. The creator should consider these standards when answering the questions outlined below. Complete, detailed answers will help with the evaluation and management of your invention and will serve to protect patent rights until an application for patent may be filed.

1. TITLE OF INVENTION:

New dual-microphone-based signal extraction algorithm that can acquire an acoustic signal faithfully in the presence of an intense noise originating from a nearby source.

2. CREATOR'S NAME(S):

Full name of those who worked on the invention (from whom the attorney can determine who may, under the law, be the creator or co-creator).

First Creator:

Name:	Albert S. Feng
Addr-work	Beckman Institute for Advanced Science and Technology, Univ. Illinois
Phone	217-244-1951
Addr-Home	1209 Wilshire Court, Champaign, IL 61821
Phone	217-359-7387

Second Creator (If any):

Name:	Charissa R. Lansing
Addr-work	Beckman Institute for Advanced Science and Technology, Univ. Illinois
Phone	217-244-2539
Addr-home	2903 Valley Brook Drive, Champaign, IL 61821
Phone	217-355-8281

Third Creator (If any):

Name:	Chen Liu
Addr-work	Beckman Institute for Advanced Science and Technology, Univ. Illinois
Phone	217-244-3067
Addr-home	2105B Orchard Street, Urbana, IL 61801
Phone	217-337-0285

Fourth Creator (If any):
Name: William D. O'Brien
Addr-work Beckman Institute for Advanced Science and Technology, Univ. Illinois
Phone 217-333-2407
Addr-home 2002 O'Donnell Drive, Champaign, IL 61821
Phone 217-359-7128

Fifth Creator (If any):
Name: Bruce C. Wheeler
Addr-work Beckman Institute for Advanced Science and Technology, Univ. Illinois
Phone 217-333-3236
Addr-home 1203 Waverly Drive, Champaign, IL 61821
Phone 217-359-0527

3. Provide a general summary of the subject matter of the invention in fifty (50) words or less. What is the purpose of the invention? Is it a new product, process, or composition of matter? A new use for or improvement to an existing product, process or composition of matter?

A new neurally-inspired dual-microphone-based acoustic signal extraction algorithm has been invented. The algorithm can be incorporated into the front-end of a hearing-aid system or an intelligent gathering system that can effectively extract an acoustic signal from a specified direction in the presence of more intense noise (up to 30 dB more intense) originating from a different source that is physically segregated from the signal source by as little as 2°.

For answers to the following questions, use remainder of sheet and attach extra sheets as necessary.

4. Provide a full and complete description of the closest known prior methods or apparatus and any disadvantages or problems of each that are solved by the present invention. As to each problem identified, provide a paragraph or so indicating:

(a) How long the problem existed and/or how long has it been appreciated to be a problem:

The problem of extracting a desired signal in the presence of intense noise is a long-standing problem in audiology and engineering. Most hearing aid devices available in the market today do not permit selective amplification of a desired signal when the signal is contaminated by noise from a nearby source (particularly if the noise is more intense than the signal). Improving the design of hearing aids, including increased directionality and noise suppression is one of the priorities for the National Institute on Deafness and other Communication Disorders of the National Institute of Health (as stated in the 1992 National Strategic Research Plan). A similar unresolved problem has existed for a very long time in the field of engineering acoustics. Namely, how can a sound that is heavily masked by intense noise from a nearby source be captured with high fidelity.

The following strategies are most commonly employed to solve this problem: use of microphone arrays (i.e., beam forming array) or a single highly directional microphone to enhance the directionality of the receiver; use of digital signal processors or an adaptive signal processor; combined use of directional microphone and adaptive beam-forming; use of personalized binaural directional hearing apparatus (to restore the directional hearing capacity of hearing impaired individuals); use of an adaptive filter; use of multi-channel compression hearing aids; use of a binaural localization scheme to locate the directions of signal and noise sources and a cross-correlation based processing or Wiener filter to obtain estimates of the desired signal; use of formant enhancement algorithms or amplification of frication (i.e., contrast enhancement algorithms).

As concluded by Lim and Oppenheim (1979), all the single-microphone speech-enhancement methods such as spectral subtraction, comb filtering, and speech-production modeling, fail to improve the intelligibility of the desired speech.

While the optimum beam-forming approach was shown to improve the intelligibility of the desired speech (Stadler and Rabinowitz, 1993; Soede, et al., 1993; Kates, 1993), this ability is greatly reduced when the speech source and the noise source are close to one another (<25°). This is due to the wide main beam which in turn resulted from the limited dimension of the microphone array. In addition, in the case of one noise source in a less reverberant environment, the noise cancellation provided by the beam-former varies with the location of the noise source with respect to the microphone array.

One two-microphone approach proposed by Bodden (1993) performs noise cancellation by means of Wiener filtering. However, since the spectra of the desired speech and the noise are not available, the estimate of the Wiener filter based on cross-correlation inevitably results in a deterioration of the fidelity of the desired speech signal (i.e., a portion of the speech signal is removed by the process of noise cancellation). This method can only suppress noise of equal intensity to that of a speech signal and the example given is for an angular separation of 30°. In addition, the approach was proposed as a

Section 4(b) redacted

(c) How the invention solves or reduces each such problem.

The approach taken is to use a neurally-inspired binaural localization scheme to locate the direction of the noise source and a proprietary mathematical algorithm to extract the desired signal at 0° azimuth.

Attach any materials, such as publications, advertisements, patents, etc., you have or that are reasonably available to you concerning the known prior methods or apparatus.

5. State the advantages of your invention over what has been done before, the problems it solves, or new applications achieved. Indicate any disadvantages or limitations and explain how they might be overcome.

The most significant advantages of the invention are the ability of the algorithm to extract a signal that is many times weaker than the noise (up to S/N of -30 dB) and the exceptional fidelity of the signal extracted. Furthermore, the required physical separation of the sources of the signal and noise is minute (as low as 2°). In other words, the directionality of the system (i.e., beam-forming characteristic) is very steep. By comparison, the existing algorithms are limited to extracting a signal when the noise is relatively much less intense (up to S/N of -6 to -10 dB) and require that the signal and noise sources be widely separated spatially (>30°). Thus, the new design represents a major improvement in terms of the spatial resolution and permits a far more extensive range of masking by noise.

6. State in general terms the purposes of the invention. Is it a new product, process, or composition of matter? Is it a new use for or improvement to an existing product, process or composition of matter?

The purpose of the invention is to significantly improve the performance of currently available hearing aid devices. The new dual-microphone-based signal extraction algorithm would allow a hearing aid user to select a desired signal in an acoustically cluttered environment for extraction and simultaneously filter out competing signals originating from a separate source.

7. Give a complete detailed description of the best model for practicing your invention with an emphasis on the new features or improvements over the known methods. Provide data or other evidence of the feasibility or operability of the invention. Attach any visual material that may be available, such as: sketches, graphs, drawings or photographs.

The algorithm was developed and tested using computer simulation. The general description of the signal acquisition system/algorithm is shown in a schematic diagram [Figure 1]. The system comprises a front end section consisting of a microphone that is connected to a frequency analyzer and a delay line; a separate analyzer and delay line are connected to each of the two microphones. With the microphones aligned toward the loudspeaker emitting the desired signal (assume this is the reference direction or 0°), the front end section is used to obtain an estimate of the azimuth (or direction) of the noise source relative to the reference. Once the direction of the noise source has been determined, a processor employing a proprietary mathematical algorithm is used to extract the acoustic signal at 0°, and at the same time removing the unwanted sound originating from the noise source.

We have simulation data to indicate that a speech signal having a peak-to-peak amplitude of 0.9 relative units [Figure 2], as emitted from a loudspeaker located at 0°, can be accurately extracted [Figure 4] when the signal is embedded in a babble noise of equal intensity [Figure 3] originating from a second loudspeaker at equal distance but located at a different azimuth (at 60° on one side of the signal loudspeaker). The algorithm performs essentially equally well [Figure 6] for an angular separation of only 2° and when the noise intensity is 30 dB above the intensity of speech signal (in this case the peak-to-peak analog amplitude of the noise is about 15.0 relative units and that of signal is 0.9 units - see

Figure 5). Although the signal as extracted by the algorithm in the latter case is somewhat noisier than that obtained under the earlier test condition, the intelligibility of the speech is excellent at least for individuals with normal hearing.

In addition to the computer simulation, tests have been carried out using actual microphones and loudspeakers in a laboratory (see item #10). These tests have provided evidence that the algorithm works effectively in a physical setting.

8. Provide details of any experiments that have been conducted, including any that have been failures, as well as all that have been successes.

See item #7 and item #10 for successes.

9. Indicate any alternate embodiments, procedures or methods of construction for the invention.

Alternate procedures and methods of construction will be evaluated as appropriate.

10. Describe the development status (concept, laboratory tested, prototype, etc.). Indicate what further development may be necessary.

The algorithm as derived from the computer simulation has been tested in the laboratory (i.e., a semi-anechoic room). One loudspeaker (L1) is used to emit speech signal and a second loudspeaker (L2) is used to produce babble noise [see Figure 1]. Two microphones, with an inter-microphone distance of 15 cm, are placed in front of L1 at a distance of about 3 feet. The L2 is placed at different azimuths but at the same distance from the mid-point of the microphones. The outputs of the microphone preamplifiers are fed into a computer (Sun Sparc 20 or a Pentium 150 Hz) and processed according to the algorithm. The algorithm requires that the L1 be placed directly ahead of the microphones and that the angular azimuth of L2 be accurately estimated. The angular estimation under the present system requires 20-30 iterations, or >100 msec. Once the estimation of spatial location is acquired, the computation leading to the extraction of the signal is fairly rapid.

To make a functional hearing aid device using the algorithm invented, we need to incorporate the algorithm into a small integrated-circuit chip. Miniaturization of the hearing aid device is thereby important for its ultimate feasibility. Also, as noted above, the computational time using Matlab is long. There is thus a need to improve its efficiency such that the device can perform on a real time basis. This performance can be readily achieved if the location of interfering noise is stationary. However, if the location of the noise source moves constantly, there is a need to streamline the algorithm for a more expeditious estimation of the angular separation between the signal and noise sources.

11. Identify the grant(s), contract(s), or other source(s) of funds contributing to the conception and development of the invention. Indicate if the invention was made as part of ones assigned duties or with the use of university facilities or services.

The funding for this research comes from in-house support from the Director of the Beckman Institute of Advanced Science and Technology, using research facilities at the Institute. With the exception of Dr. Chen Liu, the invention was not part of assigned duties of members of the research team.

12. If work on the invention is to be continued, indicate the sources of funding and the nature of the work.

This work will be continued in order to refine and improve the capability and performance of the hearing aid technology. For this, a partial support in the form of an institutional postdoctoral fellowship has been obtained for Dr. Chen Liu. Additional support will be sought from external funding agencies.

13. Give names and date of any publication(s) or abstract(s), oral or written, as well as any proposed publications which mention or describe the invention. Separate general publications from those which disclose the critical elements of the invention.

None of the results have been published or reported in any scientific meeting. A manuscript describing the design and the performance of this algorithm will be written up in the near future in preparation for submission for publication in a scientific journal.

14. Give chronology of principal events in conception and development:

(a) Earliest conception date (Is there substantiating evidence, such as a notebook or a witness?)

The conception of the idea for the project was made during the . Several members of Core Tech (Dr. Nelson Levy [Chief Executive Officer] and Dr. Eric Coles [Vice President]) are witnesses of the discussion of the plan (record is available in the form of a letter from Dr. Levy dated January 31, 1994). The formulation of the research team took place in August of 1994. There is a written record to this effect in the form of a letter to Dr. Jiri Jonas (Director, Beckman Institute of Advanced Science and Technology) requesting seed money for this project.

(b) Date of disclosure (oral or written) to other persons and names of such persons.

A written disclosure was first made to Dr. Jiri Jonas (Director, Beckman Institute of Advanced Science and Technology) on February 9, 1996.

(c) First written record and availability of such records.

Initial written record was made in September of 1995 following the joining of Dr. Chen Liu to our research team. The records are available as files on the Sparc-20 workstation, and as hard copy records in the laboratory.

(d) Dates and results of first test of invention and first successful test.

First positive results were obtained in early 1996 and the results were reported to Dr. Jiri Jonas (Director, Beckman Institute of Advanced Science and Technology) on February 9, 1996.

15. If others are known to have tried for the function or to have achieved the results of the invention and failed, describe and attach full particulars of those efforts.

None to the best of our knowledge.

Section 16 Redacted

17. If the invention or products made in accord with the invention may have been publicly disclosed, sold, offered for sale or licensed or used, then provide the dates of the first of any acts that, as the invention or products made by the invention, might be argued to be a:

(a) Publication--first date and place: Not applicable

(b) Offer for Sale--first date and place: Not applicable

(c) Offer to License--first date and place: Not applicable

(d) Public or commercial use--first date and place: Not applicable

(e) Non-laboratory or non-secret experimental use--first date and place: Not applicable

First Creator Albert S. Feng

Date: May 8, 1996

18. Each page of the disclosure should be signed and dated by the creator(s), and then read and signed by a witness, or witnesses who understands it, using the following statement:

DISTRIBUTION Prepare and distribute 5 copies of the completed Invention Disclosure Form as follows:

1 copy for your file
1 copy to Unit Executive Officer
1 original and 2 copies to the Research and Technology Management Office, 417 Swanlund Admin. Building, 601 E. John Street, Champaign, IL 61820 (MC-304)

FIGURE CAPTIONS

Figure 1 Schematic diagrams of the experimental setup. A. Top view of the spatial arrangement of the components. B. Diagram of the computational scheme.

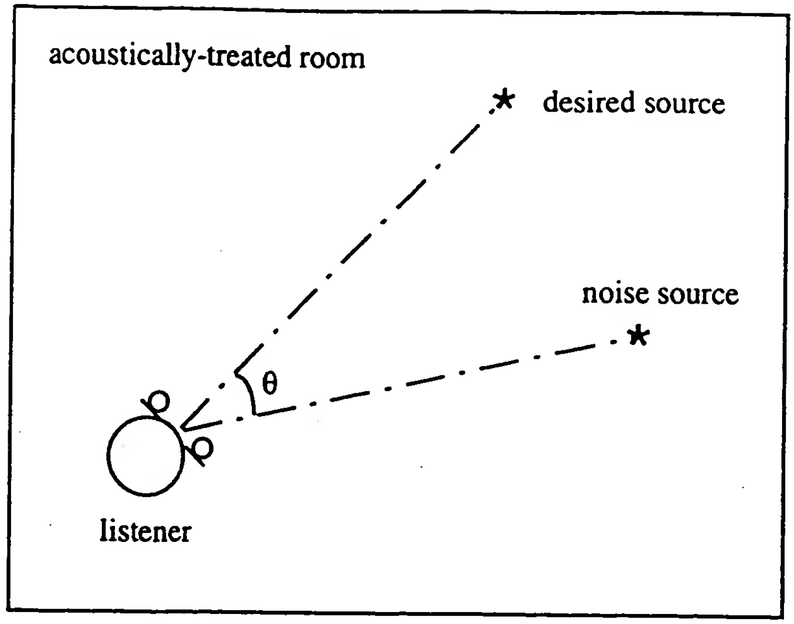
Figure 2 Speech signal in the form of a sentence (2 sec long) emitted from a loudspeaker at 0°.

Figure 3 Speech signal embedded in babble noise of equal intensity (0 dB SNR) originated from a neighboring loudspeaker which is spatially separated (60° away) from the signal source is highly noisy.

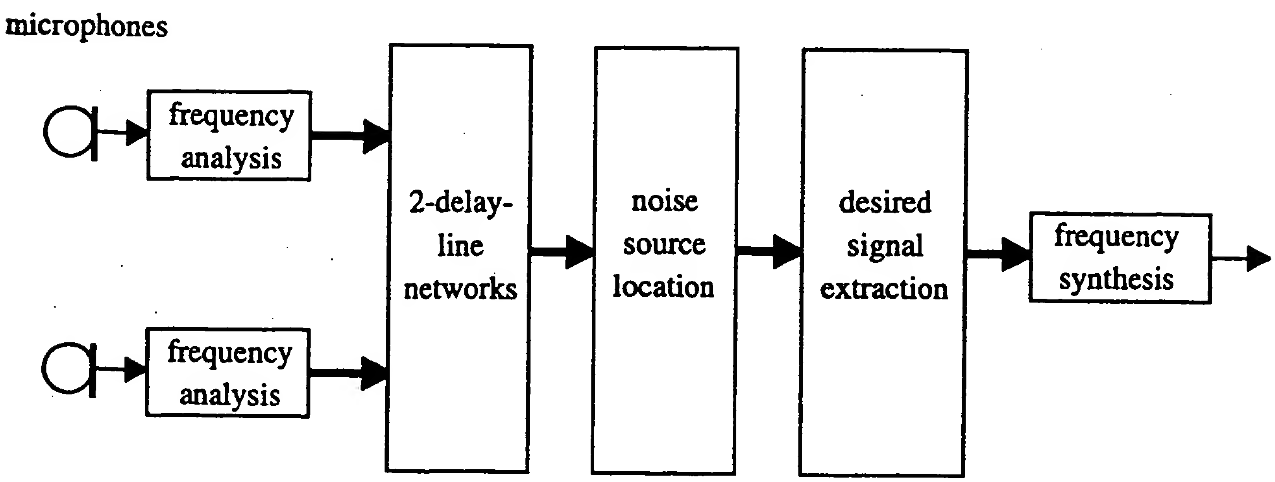
Figure 4 Signal recovered by the neural algorithm is essentially perfect - as clean as the original speech signal (shown in Figure 2).

Figure 5 Speech signal embedded in babble noise of high intensity (-30 dB SNR) originated from a neighboring loudspeaker which is only 2° away from the signal source is extremely noisy. Note that the scale in the ordinate is 20x higher than in Figure 1.

Figure 6 Signal retrieved by the neural algorithm is excellent, albeit slightly noisier than the original speech signal. The signal is clearly intelligible (it sounds like having a weakly audible hiss added to the speech signal shown in Figure 2)



A. Top view of the spatial arrangement of the components.



B. Diagram of the computational scheme.

Figure 1

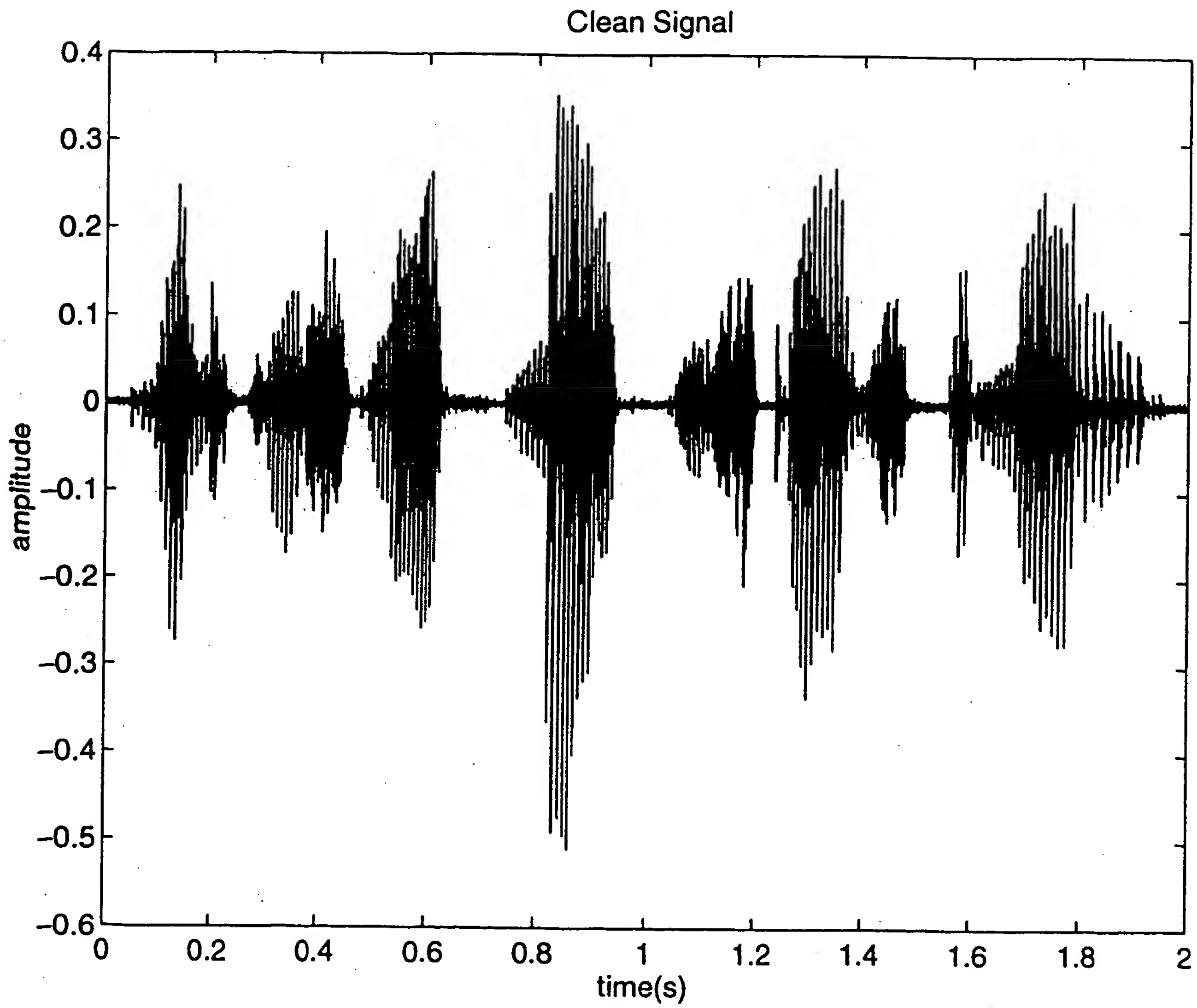

Clean Signal
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

Figure 2

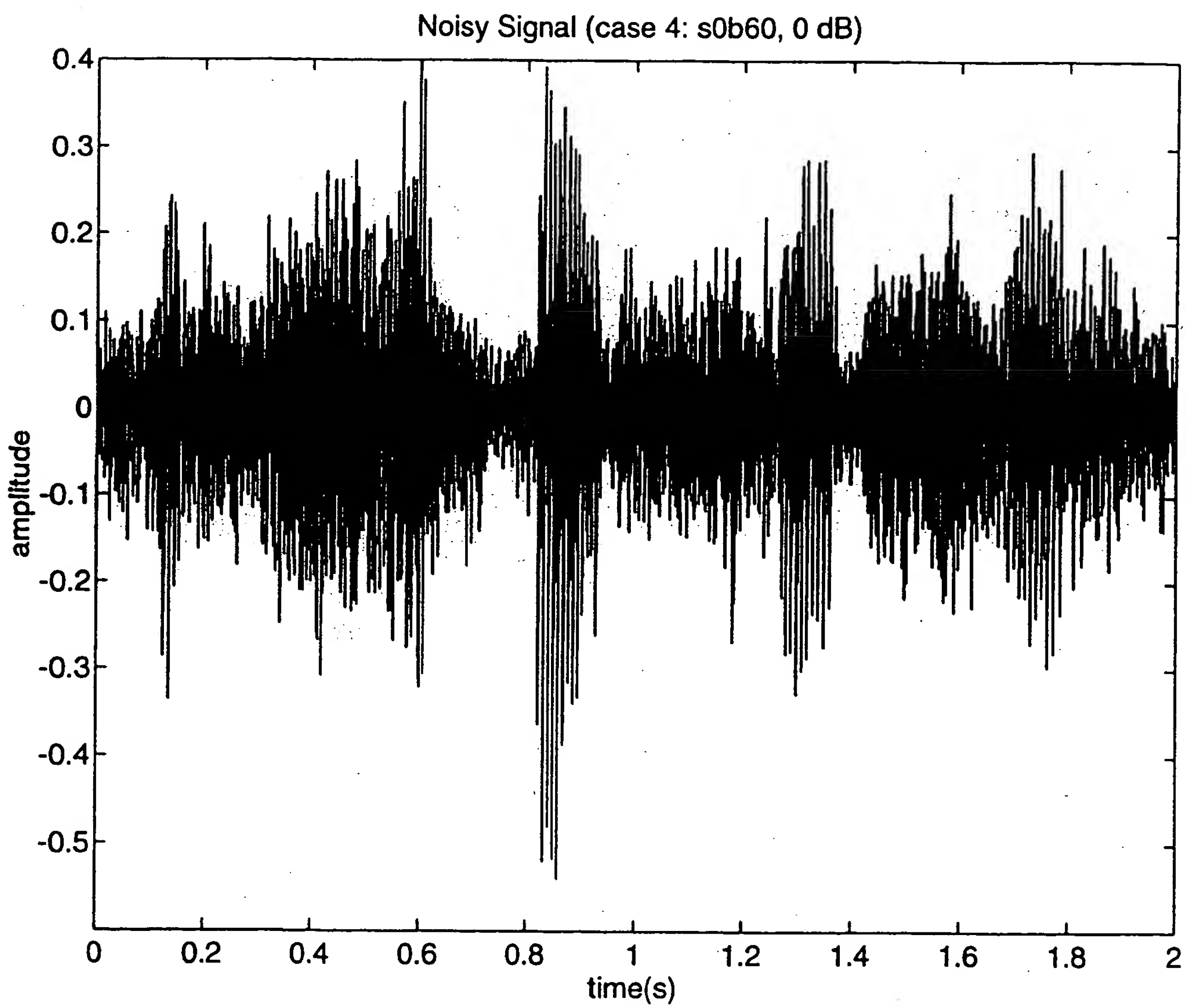

Noisy Signal (case 4: s0b60, 0 dB)
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

Figure 3

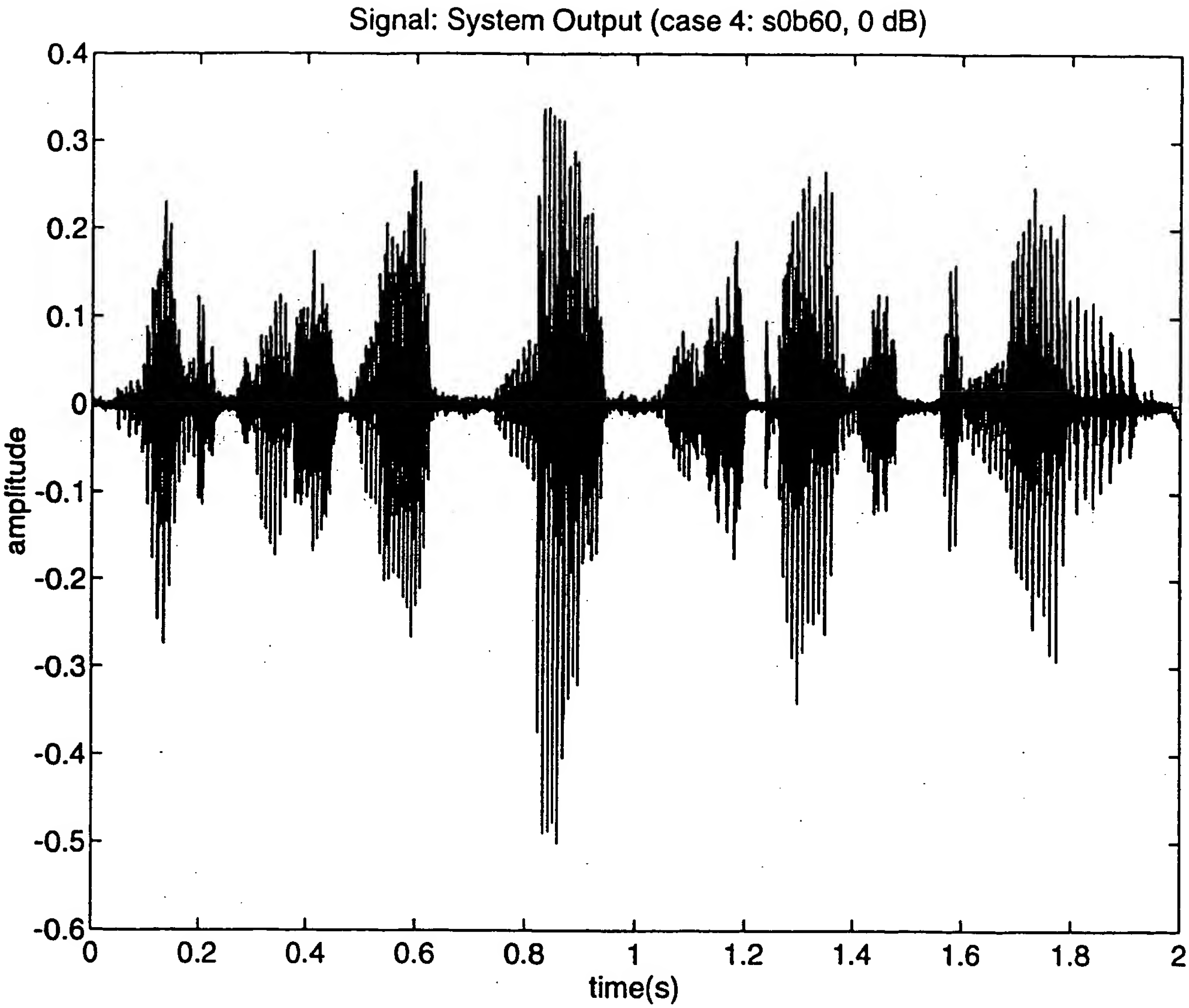

Signal: System Output (case 4: s0b60, 0 dB)
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

Figure 4

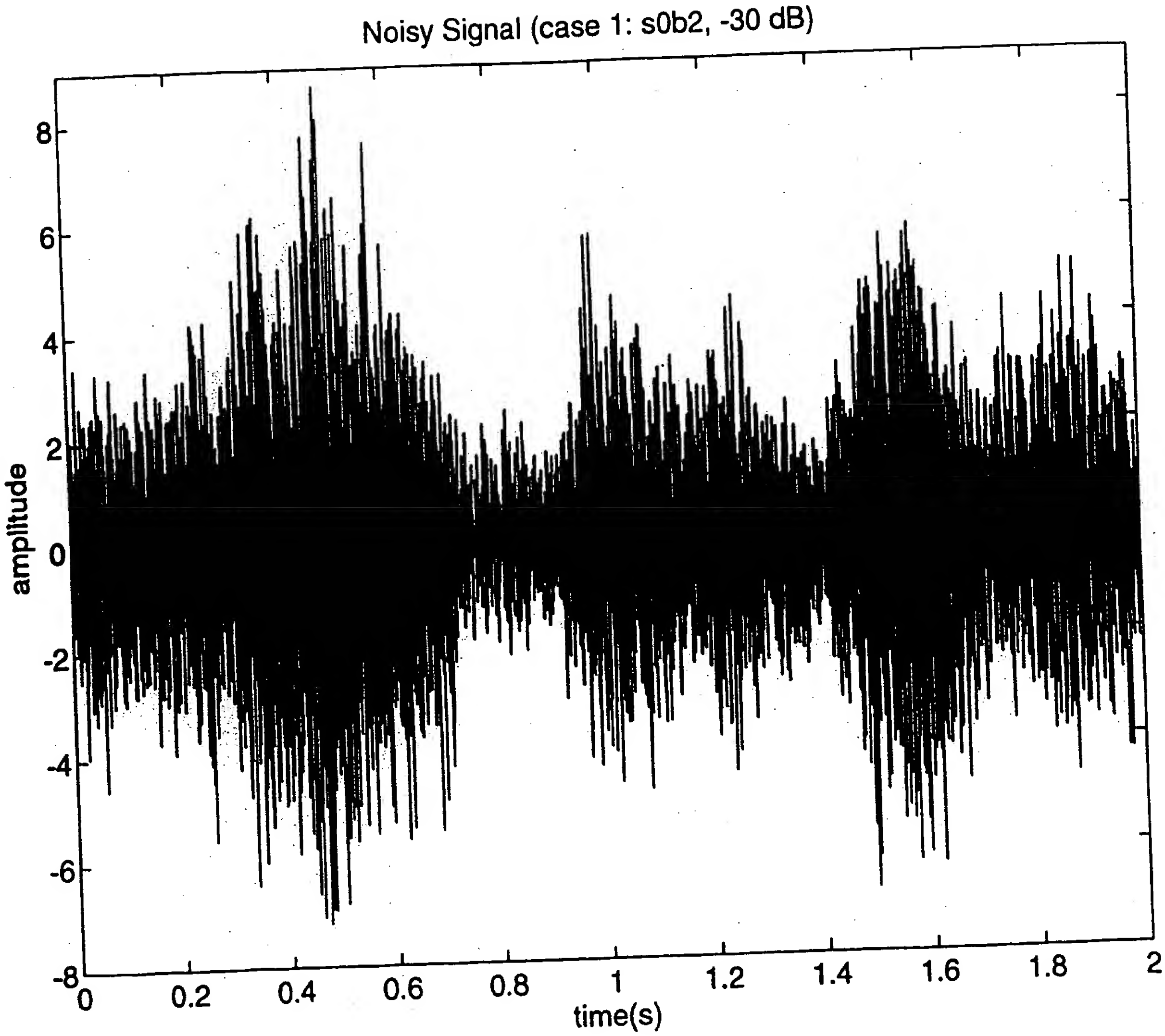
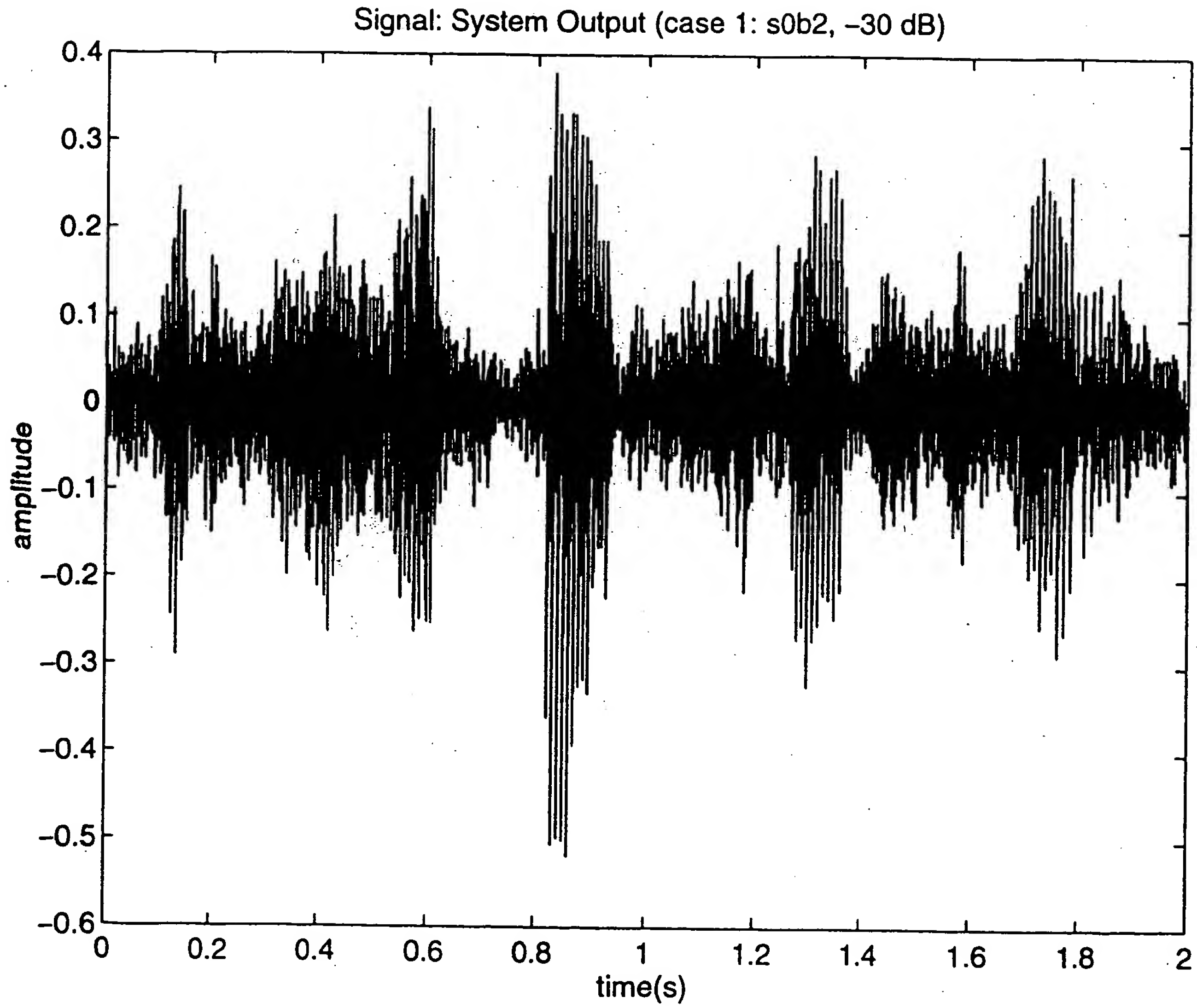

Noisy Signal (case 1: s0b2, -30 dB)
amplitude
time(s)
8
6
4
2
0
-2
-4
-6
-8
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Figure 5



Signal: System Output (case 1: s0b2, −30 dB)
0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
amplitude
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
time(s)

Figure 6

APPENDIX

Described herein is the information processing scheme of a Hearing Aid System which is designed to extract a desired signal, emitted at one location, embedded in noise emitted at a second location. In usual practice the desired signal would be the speech of a person directly in front of the listener, while the noise could be any interfering sound including the speech of another individual. This description includes the algorithmic components beginning with the digitization of the sound at two locations and ending with the reconstruction of the desired speech signal.

(i) This Hearing Aid System utilizes two identical microphones as receivers (denoted #1 in Figure 1A); these are separated from one another by a fixed distance and positioned in such a way that the sound source emitting the desired signal is directly ahead (i.e., *on-axis*) of the mid-point of the microphone pair. The other sound source is *off-axis* at a different sound direction. Signals as picked up by the two microphones, $x_{Lp}(t)$ and $x_{Rp}(t)$, are fed to two processing channels. Here L and R denote left and right processing channels, respectively, and p denotes the time frame of the short-term spectral analysis (see below).

(ii) The signals as picked up by the microphones are filtered to prevent aliasing and digitized by analog-to-digital converters (#2 in Figure 1A). The digital versions of the signals are designated as $x_{Lp}(k)$ and $x_{Rp}(k)$, respectively, where the index k refers to the discrete time at which the samples are taken.

(iii) The digitized signals are transformed into the frequency domain by means of short-term spectral analysis (#3 in Figure 1A) across the entire audible frequency range. This process can be realized in practice by the Discrete Fourier Transform (DFT), or by means of a filter bank. The transformed outputs are complex signal amplitudes, $X_{Lp}(m)$ and $X_{Rp}(m)$, evaluated at discrete frequencies f_m, (m=1, ..., M).

(iv) For each frequency, the complex signal amplitudes from the two channels are fed into a pair of delay-lines (#4 in Figure 1A), each of which has an even number N of delay units (and stores N+1 values including the current value). Conceptually, the signals from the left microphone are propagated from left to right, while those from the right microphone are propagated from right to left. An acoustic signal originating from any direction will be in phase at one specific locus along the length of the dual delay-line. The values of the time delays are assigned *a priori* such that the acoustic space in front of the two microphones is

divided uniformly into N+1 directions in its azimuth and each sound azimuth is uniquely mapped to one location along the dual delay-line.

(v) Each signal pair $X_{Lp}(m)$ and $X_{Rp}(m)$ in the dual delay-line is input to a computational unit (#5 in Figure 1A and 1B), which performs the following computational algorithm:

$$X_p^{(i)}(m) = \frac{X_{Lp}^{(i)}(m) - X_{Rp}^{(i)}(m)}{\exp\left[-j2\pi(\tau_i + \cdots + \tau_{\frac{N}{2}})f_m\right] - \exp\left[j2\pi(\tau_{\frac{N}{2}+1} + \cdots + \tau_{N-i+1})f_m\right]} \quad \text{for } i \le \frac{N}{2}$$

or

$$X_p^{(i)}(m) = \frac{X_{Lp}^{(i)}(m) - X_{Rp}^{(i)}(m)}{\exp\left[j2\pi(\tau_{\frac{N}{2}+1} + \cdots + \tau_{i-1})f_m\right] - \exp\left[-j2\pi(\tau_{N-i+2} + \cdots + \tau_{\frac{N}{2}+1})f_m\right]} \quad \text{for } i > \frac{N}{2}+1$$

The locations on the dual delay-lines are both indexed from left (i=1) to right (i=N+1). The τ_i are the values of the time delays.

(vi) At the outputs of the computational units, an average of the energy is derived for each location (i) across the frequency bins m = 1, ..., M (#6 in Figure 1A):

$$\overline{\overline{X_p^{(i)}}} = \frac{1}{M}\sum_{m=1}^{M}\left|X_p^{(i)}(m)\right|^2$$

(vii) A time average of the output of $\overline{\overline{X_p^{(i)}}}$, over P most recent spectral-analysis time frames, is then computed as follows:

$$\overline{\overline{X^{(i)}}} = \sum_{p=1}^{P}\gamma_p \overline{\overline{X_p^{(i)}}}$$

where γ_p are empirically determined weighting factors. The noise source localization unit (#7 in Figure 1A) makes an estimate of the azimuth of the noise source by finding the location of the global minimum $\overline{\overline{X^{(i_{noise})}}}$ of $\overline{\overline{X^{(i)}}}$ along the dual-delay-line (output of #6 in Figure 1A):

$$\overline{\overline{X^{(i_{noise})}}} = \min[\overline{\overline{X^{(i)}}}]$$

(viii) The noise source localization unit controls the input selection (#8 in Figure 1A) of the speech reconstruction module (#9 in Figure 1A), to pinpoint the i_{noise}-th column of the output matrix of processing unit (#5 in Figure 1A), $X_p^{(i_{noise})}(m)$, for which the noise is maximally canceled. Thus, $X_p^{(i_{noise})}(m)$ provides the best estimate, symbolized by $\tilde{S}_p(m)$, of the spectrum $S_p(m)$ of the desired signal.

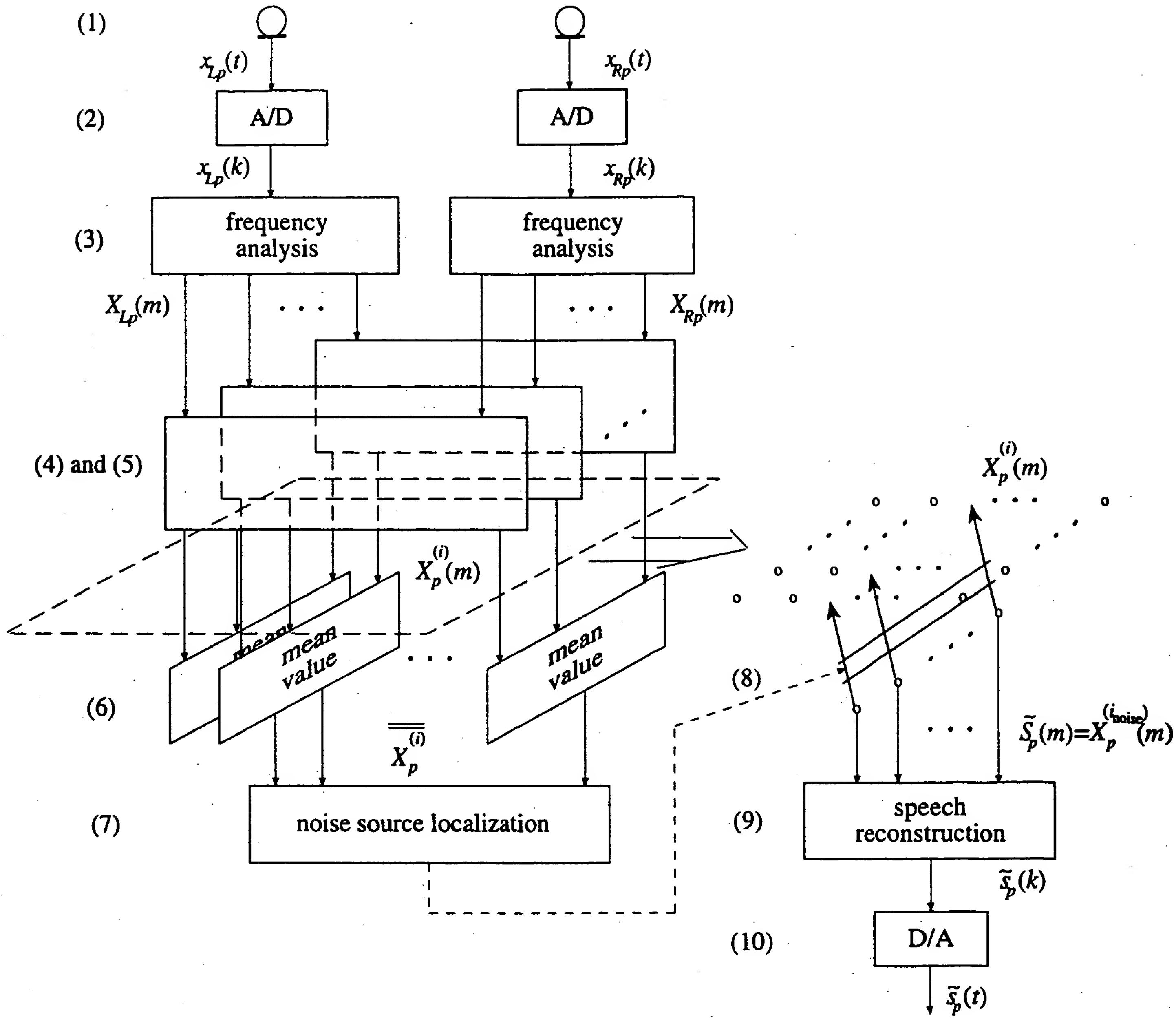
$$\tilde{S}_p(m) = X_p^{(i_{noise})}(m)$$

(ix) The speech reconstruction module (#9 in Figure 1A) converts the signal spectrum estimate $\tilde{S}_p(m)$ to the time domain:

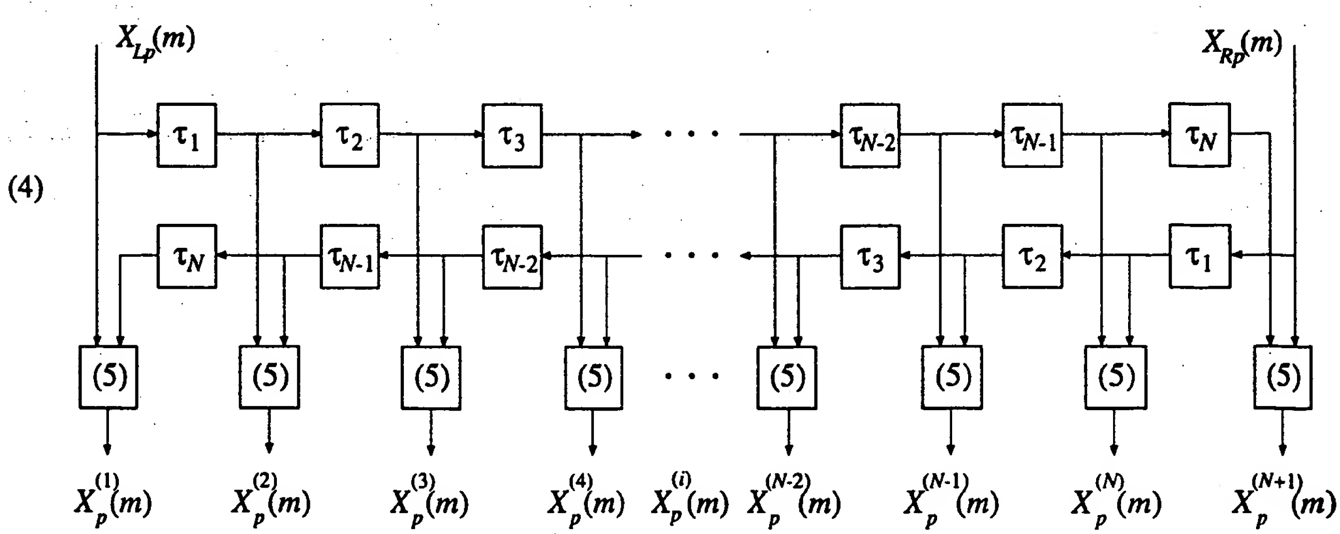
$$\tilde{s}_p(k) \Leftrightarrow \tilde{S}_p(m)$$

(x) The digital version of the speech signal $\tilde{s}_p(k)$ is then converted into its analog form [$\tilde{s}_p(t)$] by means of a digital-to-analog converter (#10 in Figure 1A).

Figure 1



(A) Block diagram of the system.



(B) Details of the 2-delay-line and accompanying calculation units.

Extraction of Signal Embedded in Noise

What follows is consideration of separation of two sound sources over a *single* frequency (denoted by ω). By means of Fourier transform, the separation of two *complex*-sound sources can be accomplished.

The source of desired signal is assumed to be straight ahead; while the interfering sound (or noise) source is off-axis. Two omnidirectional microphones are employed to detect the signal and noise. The outputs of these microphones are filtered, digitized, Fourier transformed, and fed into a pair of delay-lines (see Appendix for the arrangement of the paired delay lines). The tapped-outputs of the two delay lines are paired as shown in Fig. 1. Given the azimuthal position of the source of the desired signal, the components of the desired signal should be in phase at the outputs ϕ_1 and ϕ_2 at the midpoint of the delay lines. This corresponds to position I on the delay line (Fig.1). Similarly, the components of the interfering sound are in phase at the outputs ϕ_3 and ϕ_4 at a different position marked II in Fig.1 (see Appendix for finding this position). The phase distance between the positions I and II is assumed to be Δ.

We denote the desired signal at position I at an arbitrary instant t as

$$A_0 e^{j(\omega t+\theta_0)} \tag{1}$$

and the interfering sound at position II as

$$A_1 e^{j(\omega t+\theta_1)}. \tag{2}$$

where θ_0 is phase angle of the desired signal at position I, and θ_1 is phase angle of the interfering sound at position II. The outputs ϕ_1 and ϕ_2 of the delay lines at the midpoint I at the same instant t are, respectively,

$$\phi_1 = (A_0 e^{j\theta_0} + A_1 e^{j(\theta_1+\Delta)})e^{j\omega t}, \tag{3}$$

and

$$\phi_2 = (A_0 e^{j\theta_0} + A_1 e^{j(\theta_1-\Delta)})e^{j\omega t}. \tag{4}$$

The outputs ϕ_3 and ϕ_4 of the delay lines at position II at the instant t are, respectively,

$$\phi_3 = (A_0 e^{j(\theta_0-\Delta)} + A_1 e^{j\theta_1})e^{j\omega t}, \tag{5}$$

and

$$\phi_4 = (A_0 e^{j(\theta_0+\Delta)} + A_1 e^{j\theta_1})e^{j\omega t}. \tag{6}$$

The factor $e^{j\omega t}$ can be ignored since it is the same not only for both the desired and the interfering sound, but also for the outputs at the different positions on the delay lines as long as the outputs are taken at the same time instant t. Thus the amplitude as well as the initial phase of the desired signal at instant t can be estimated by

$$A_0 e^{j\theta_0} = \frac{\phi_4 - \phi_3}{e^{j\Delta} - e^{-j\Delta}} . \tag{7}$$

This estimation can be obtained provided the position of the source of interfering sound (Δ) is known or can be accurately determined.

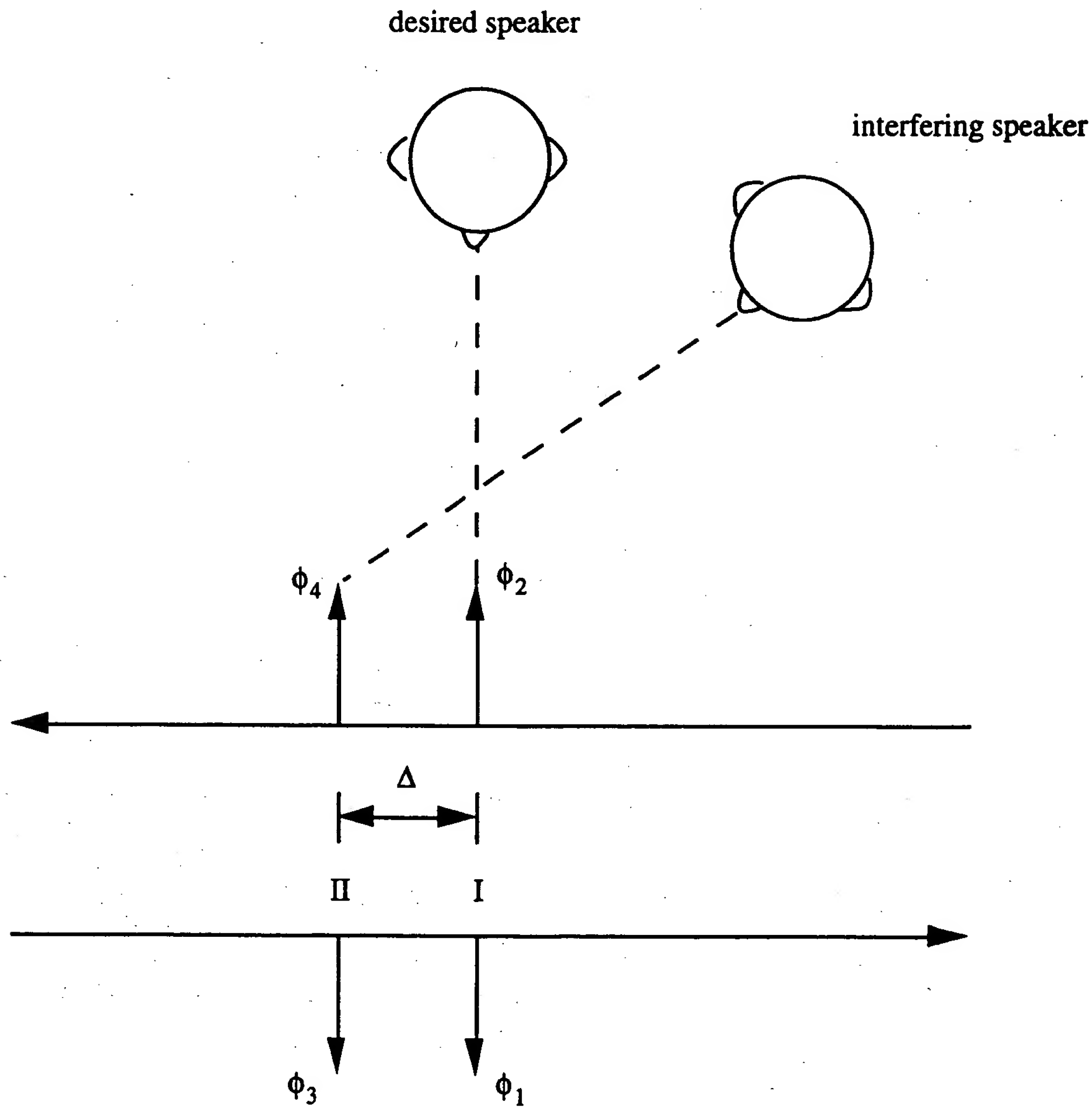


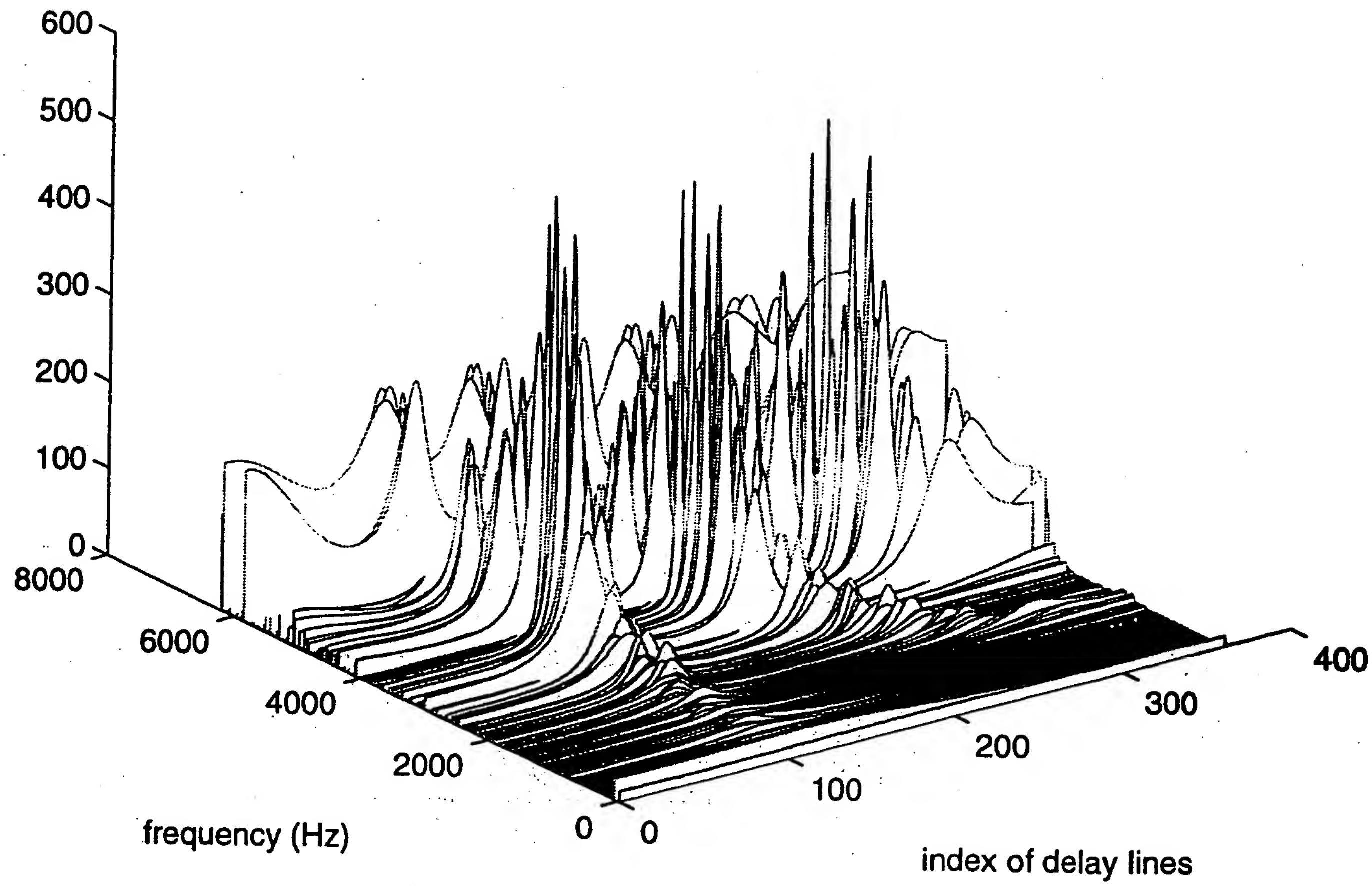
Fig. 1. Schematic diagram of signal extraction algorithm.

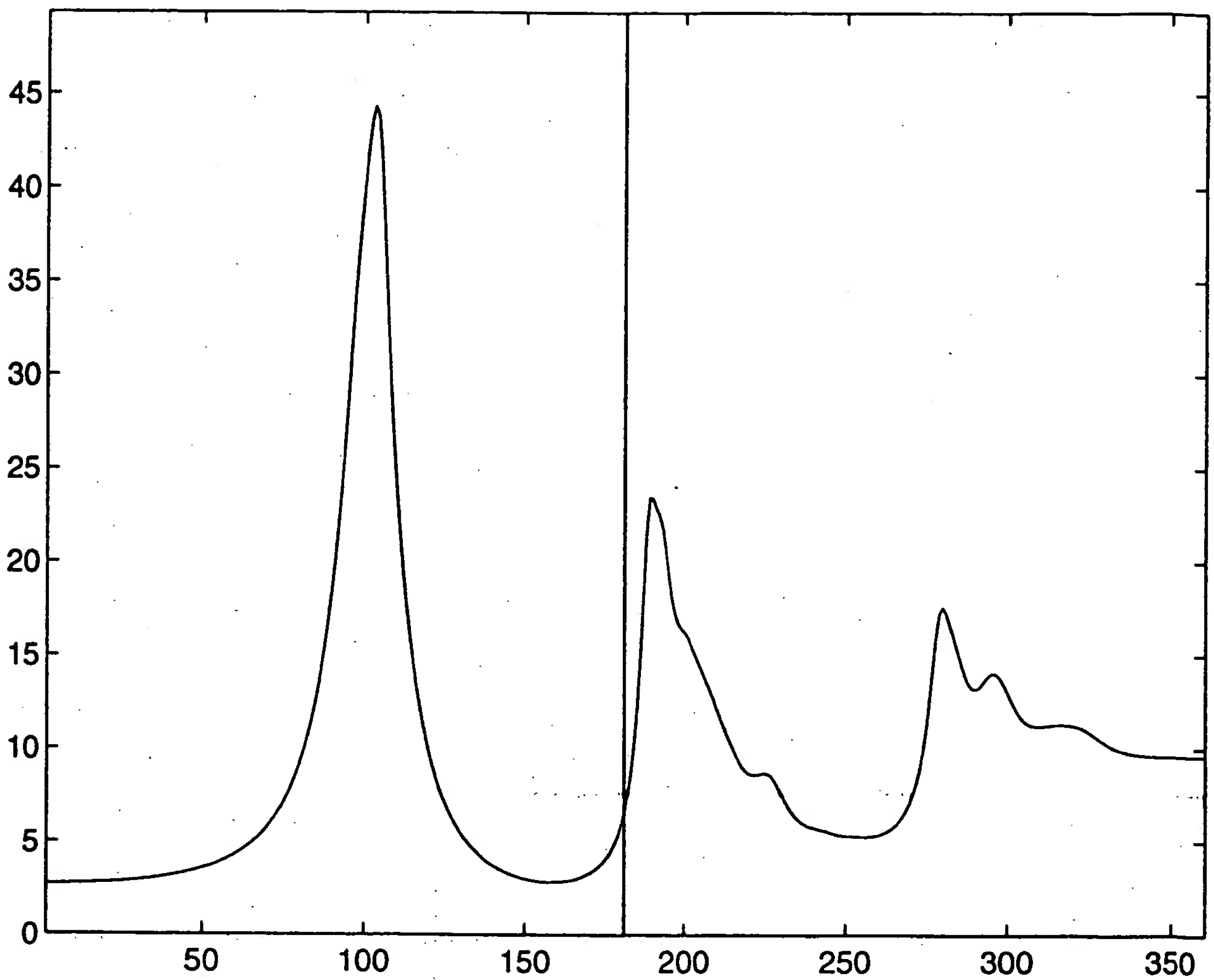
Derivation of global minima for determining the direction(s) of sound source(s)

Two figures are shown here to illustrate the concept of global minima (or maxima) that is used for determining the direction(s) of sound source(s). These figures are derived from a new algorithm which represents an important improvement over the original algorithm. The improvement is in the capacity to determine the directions of both the desired signal and the interfering sound. With this improvement, we essentially remove the needs to align the microphone pair toward the desired signal. Instead, the direction of the desired signal can be any arbitrary direction close to the *on-axis*. In the situation where there are two sound sources, the sound originating from a location closer to the *on-axis* will be treated as the desired signal whereas the sound originating from *off-axis* will be treated as noise (or interfering sound), or vice versa.

Figure 1 shows the 3-dimensional plot of values of the equivalent of inverse of energy [as used in the original algorithm], from the 18th spectral-analysis time frame. The x-axis is the index (i) of the position on the dual delay-line, y-axis is the index (m) of each frequency bin, and z-axis indicates the energy of each frequency component at each position on the dual delay-line. The locations of the peaks of local maxima (see Figure 2 for global maxima) on the dual-delay-line indicate the azimuthal positions of *both* the desired signal *and* the interfering sound.

Figure 2 displays the average of the inverse of the energy derived for each location (i) across the frequency bins (see Figure 1). The peaks represent the global maxima which correspond to the locations of the sound sources in free-field. The third global maxima on the far right is an artifact that, for the practical purpose, can be ignored.


600
500
400
300
200
100
0
8000
6000
4000
2000
0
frequency (Hz)
0
100
200
300
400
index of delay lines



Index of delay lines

Date: Tue, 25 Oct 1994 10:50:11 -0500
X-Sender: wheeler@ux1.cso.uiuc.edu
Content-Type: text/plain; charset="us-ascii"
To: feng@ux1.cso.uiuc.edu, Charissa Lansing <crl@uiuc.edu>, yxz@enterprise.ifp.uiuc.edu
From: bwheeler@uiuc.edu
Subject: some notes

Al, Charissa, and Yunxin,

This is only a start, but I thought I'd better get you something rather than nothing.

Bruce

Intelligent Hearing Aid
Notes and Outline
Bruce Wheeler
10/19/94

Preamble:

Our current working rationale is as follows:

While there has been tremendous progress in the signal processing associated with hearing aids, including filters with gains which are programmable, adaptive, and frequency band specific, there are two types of information which are underexploited and which should be the basis, either singly or in combination, for a new generation of devices. These are:
1. spatial information, gained from multiple microphone receivers, especially with directional selectivity, with new strategies for optimal physical placement and appropriate signal processing
2. speech unit specific information, which can be detected by digital pattern recognition and used to control signal gain in order to enhance the parts of speech most important for improved intelligibility by the hearing impaired subjects

These observations have support in the literature, which we are actively exploring, and lead to a plan for research spanning:
1. the physiological bases for hearing deficits, especially as it relates to aging,
2. the psychoacoustical documentation of these deficits,
3. the evaluation of prescriptive remedies which include both new hearing aids and strategies for use with visual information (e.g. lip reading),
4. a combination of strategies for microphone placement and signal processing for directionality and noise cancellation
5. signal processing for speech unit recognition coupled to enhanced filtering
6. an architecture for a hearing aid chip, which allows flexible development of both spatial and speech component dependent amplification
7. microminiaturization for ultimate feasibility of an in the ear or in the canal hearing aid.

Background:
1. State of the art in analog and digital hearing aids
2. Understanding of the physiology of hearing deficits, especially as related to aging
3. Review of what is known about increasing intelligibility of speech with

various signal manipulations

Bruce C. Wheeler, Associate Professor
Electrical and Computer Engineering Department,
Neuroscience Program, and Bioengineering Faculty
Beckman Institute, University of Illinois at Urbana-Champaign
405 N. Mathews, Urbana IL 61801;
217-333-3236; FAX: 217-244-5180 or 217-244-8371
email: bwheeler@uiuc.edu

University of Illinois
at Urbana-Champaign

Beckman Institute
405 North Mathews Avenue
Urbana, IL 61801

December 27, 1994

TO: Bruce Wheeler
Charissa Lansing
Yunxin Zhao

FROM: Al Feng

RE: Funding from BI

Good news! Margarita Ham just informed me that our request for equipment has been approved for the full amount. I should be getting the official notification from her shortly.

We can now order all the equipment on our list. I have already told Bob Penka that we can contribute $2K toward the software license. We should order the rest as soon as we can. We do not have the experimental room assigned to us yet but I hope to hear from Sarah shortly.

Date: Fri, 27 Jan 1995 14:06:01 -0600
X-Sender: penka@ux1.cso.uiuc.edu
Mime-Version: 1.0
To: morgan@cogsci.uiuc.edu, feng@ux1.cso.uiuc.edu, rbargar@ncsa.uiuc.edu,
yxz@ifp.uiuc.edu.r-sousa
From: r-penka@uiuc.edu
Subject: ESPS/waves+/HTK train finally leaves the station!

I have started the paperwork to order for ESPS/waves+/HTK. Before Entropics can ship the software I must tell them:
1) the platforms we have and the media we need
2) the number of licenses we need
3) the identify of the support person in each participating unit who will be entitled to contact Entropics for software support. Entropics expects one support person per contributing unit , but this might be negotiable. I named the following units as contributors:
Beckman Institute
College of Liberal Arts and Sciences
Department of Electrical and Computer Engineering
NCSA
Department of Spanish, Italian and Portugese
Department of Linguistics
Everyone within these units will be licensed to use ESPS/waves+/HTK.

To enable me to supply this information to Entropics would you please tell me:
A) the name of the support person for the unit you represent. This person will have access to Entropics techincal support. To eliminate ambiguity for those of you in Beckman, assume the following pairings of names and units:
Bargar - NCSA
Zhao - Computer and Electrical Engineering
Feng - Beckman Institute
Morgan - Linguistics
Sousa - Spanish, Italian and Portugese

We could name a sixth person (for LAS) but I'm at a loss for how to identify that person. Jerry Morgan, could you suggest someone?

B) The number of licenses your (as defined above) unit will need.

C) the media your unit can handle and the platforms on which it will run ESPS/waves+/HTK. Ken Nelson of Entropics tells me that ESPS/waves+/HTK is available only for:
Sun/4 or SPARCstations running Solaris or Sun/OS
SGI Indy and Indigo workstations running system versions 4.05F - 5.2
HP 9000 running HP/UX
DEC Alpha workstations running OSF/1
DECstation 3100 or 5100 running Ultrix

Media available: 4-mm tape
8-mm tape
QIC tape cartridge (1/4 inch tape)

Thank you
Robert Penka

X-Sender: penka@ux1.cso.uiuc.edu
Mime-Version: 1.0
Date: Tue, 14 Feb 1995 15:20:31 -0600
To: l-haken@uiuc.edu, feng@ux1.cso.uiuc.edu, rbargar@ncsa.uiuc.edu
From: r-penka@uiuc.edu
Subject: ESPS/waves+/HTK

Before Entropic software will ship ESPS/waves+/HTK I must supply them with the information requested below. Entropic asks that I tell them everything we need up front so that they can include everything in a single shipment.

1) Please tell me the platforms on which you plan to run the software.
Here are the available platforms:
SUN workstations (SUN/OS)
SUN workstations (Solaris)
SGI Indy and Indigo workstations (system versions 4.05 through 5.2)
DEC Alpha workstations (OSF/1)
HP 9000/700 (HP/UX)
DECstation (Ultrix)

2) For each platform you will need, please identify the media you require.
The available formats are:
4mm DAT
8mm exabyte
1/4" cartridge

3) Please identify the number of servers you will have and the number of licenses each server will manage. E.g. (license server #1, 15 simultaneous users), (license server #2, 25 simultaneous users), ...

4) Please identify a technical support person for your use of ESPS/waves+/HTK. Entropic technical support will accept incoming calls from this person. [The contract permits us to name one person (must be an employee, not a student) per contributing department. It's quite likely that they will relax this rule and permit us to name one person per site. Give the names of the persons you would like named but, just in case, identify the primary person you want named if Entropic will give you only one slot.

I would appreciate a prompt reply. Entropic has received the signed license agreement and the Purchase Order. They will ship as soon as I give them this information.

August 25, 1994

Mr. Chen Liu
Department of Biomedical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 32000
Israel

Dear Mr. Chen:

I read with great interest your letter of July 21, 1994 in which you inquired about a postdoctoral position in my laboratory. Your research interest matches mine almost perfectly. Students and postdocs in my laboratory are presently engaged in physiological studies aiming toward the understanding of physiological mechanisms that underlie coherent perception of sounds in "noisy" environment, i.e., similar to a "cocktail party". Several colleagues of mine from the Beckman Institute for the Advanced Science and Technology (my home base) and I have also begun collaborative work to pursue research on designing intelligent hearing aid devices. I therefore think you would fit in well with our research programs. Further, I am in a position to offer you postdoctoral salary, at least for your first year.

In looking over your CV, I notice that you did not list the names of references. As is customary for postdoctoral applications, however, I would appreciate receiving three letters of recommendation from professors who can provide frank assessments of your scientific qualifications as well as personal characters. I would also welcome their comments regarding your productivity as a graduate student or research associate. Finally, to further assist me in the evaluation, can you please send me reprints of your past publications and copy of manuscript(s) that describes your present dissertation work?

I look forward to receiving these materials in the mail, or through fax (217-244-5180).

Sincerely,

Albert S. Feng
Professor of Physiology & Bioengineering
Tel: 217-333-1734
Email: FENG@UX1.CSO.UIUC.EDU

הטכניון - מכון טכנולוגי לישראל
המחלקה להנדסה ביו-רפואית
המכון למדעי הנדסה רפואית וביולוגית ע"ש יוליוס סילבר
TECHNION — ISRAEL INSTITUTE OF TECHNOLOGY
Department of Bio-Medical Engineering
The Julius Silver Institute of Bio-Medical Engineering Sciences

February 19, 1995

Professor Albert S. Feng
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
Department of Physiology and Biophysics
524 Burrill Hall
407 South Goodwin Avenue
Urbana, IL 61801 USA

Dear Professor Feng,

My passport has just been replaced. The information in the new passport is given below.

Family Name:	LIU
First Name:	Chen
Birth Date & Place:	3 October 1964, Tianjin, P.R. China
Marital Status:	Married
Citizenship & Country of Legal Permanent Residence:	P.R. China
Dependent:	Qing Qi - Wife
*Passport #:	2806748
*Issuing Place:	Tel-Aviv, The Embassy of P.R.China in Israel
*Expiration Date:	February 14, 2000

Only the items labeled with asterisks are changed. My wife's personal information remains the same as I sent to you on October 19, 1994.

Family Name:	QI
First Name:	Qing
Birth Date & Place:	24 January 1969, Yunnan, P.R. China
Marital Status:	Married
Citizenship & Country of Legal Permanent Residence:	P.R. China
Passport #:	2186806
Issuing Place:	Tianjin, P.R. China
Expiration Date:	February 22, 1998

In case you need further information, please contact me either by email (Liu@biomed.technion.ac.il) or by fax (00972-4-234 131). I am sorry for this change and the inconvenience incurred.

With best regards,

Chen Liu

Chen Liu

Technion City, Haifa 32 000, Israel. Tel.: (04) 239431 ,294130 :קרית הטכניון, חיפה 32 000, טל
TELEX : 46406 TECON IL ; FAX.: 972-4-234131

ae

To: Chen Liu <liu@biomed.technion.ac.il>
From: feng@ux1.cso.uiuc.edu
Subject: Re: schedule change
Cc: ASF

Dear Mr. Liu:

Happy New Year to you too!

I can understand the need for a schedule change. It happens quite often for our students. A delay of one month does not pose a problem (but if it goes far beyond one month we may have to reconsider our offer).

With best wishes.

Albert Feng

>Dear Professor Feng,
>
>First of all, I would like to take this opportunity to wish you a happy
>Chinese Spring Festival.
>
>I am writing to you due to a quite unexpected change in my plans. My
>superviser, Prof. S. Sideman, has changed his Sabbatical schedule and,
>despite my request otherwise, he has postponed my final examination.
>Therefore, I will only be able to come to the US around late August, 1995.
>I am very sorry for this development.
>
>I hope that this will not unduly affect my plans to come to the US for my
>post-doctoral work.
>
>Also, for your information, my Chinese passport is being renewed. I shall
>advise Ms. M. Ham of my new passport number as soon as I receive it.
>
>I am very sorry for the inconvenience that this change in my exam schedule
>may incur, but hope that you understand and look forward to beginning my
>work with you soon. Please let me know how this schedule looks to you,
>and I thank you in advance for your understanding in the matter.
>
>With best regards,
>Chen Liu

November 29, 1995

TO: Hearing Aid Research Faculty

FROM: Al Feng

RE: Draft for proposal for the Critical Research Initiative

Attached please find a draft for proposal that is to be submitted to the CRI for obtaining support for our project.

Specific Aims (Draft)

Over 28 million Americans have hearing impairments that restrict their ability to communicate. Of these, about 5 million employ hearing aid devices to improve their ability to hear and to communicate. A survey of hearing impaired subjects whowear hearing aid devices indicates that, in spite of the advances in the electronic technology, only 58% of them found the current generation of hearing aid devices to be adequate for their needs. The limited satisfaction of hearing aids is in part attributed to the wider range of acoustic environments in which the hearing impaired subjects dwell, a situation that was made possible by successful miniaturization of hearing aid devices. The limitation comes from the fact that the hearing aid apparatus usually amplifies all sounds including the desired signal as well as the competing background noise (or unwanted signals).

A collaborative research team at the Beckman Institute has launched an interdisciplinary effort to design and construct intelligent hearing aid devices that selectively amplify speech sound (i.e., the signal) embedded in the background noise originating from different sectors of auditory space. Selective speech processing in the normal hearing subject depends on the ability of the nervous system to focus on sounds emanating from a narrow sector of auditory space. The desired speech signal can be successfully deciphered even in the presence of intense background noise so long as the origins of the speech signal and noise are spatially separated.

The fundamental approach is to use neurally inspired scheme to develop a highly directional signal acquisition system which can effectively acquire desired signal originating from a small sector of auditory space such that interfering sounds emanating from other sectors of auditory space are deemphasized. Once the speech signal is captured, it would be appropriately amplified and its speech content enhanced, using the state-of-the-art speech enhancement algorithms, to improve the speech intelligibility. This research has two specific aims.

Aim #1 is to test the hypothesis that an effective signal acquisition (SA) system for acquiring a signal in noisy background can be achieved by a neurally inspired system. The real-time SA system shall consist of strategically-placed microphones which can be directed into a segment of auditory space for picking up speech sounds originating from that sector of auditory space. To achieve high directionality of the microphone system, signals picked up by the two microphones shall be processed using neurally inspired algorithms known to be highly effective. For this, the desired signal will be decomposed into its Fourier components, processed, and recomposed to restore to its original form.

Approach:

1. Employ different placements of microphones and different neural algorithms to optimize signal acquisition and noise cancellation - (PIs: Liu, Feng, O'Brien, Wheeler, and TBA).
2. Evaluation of the effectiveness of the SA system in normal listening subjects - (PIs: Bilger, Gupta, and TBA).

Aim #2 is to test the hypothesis that an intelligent hearing aid (IHA) system for effective speech recognition in noisy background can be reconstructed by combining the multiple-microphone based SA system (described in Aim #1) with: (a) speech enhancement algorithms (i.e., speech-unit amplification and/or filtering algorithms), (b) visually-based lip reading mechanism.

Approach:

1. Test different speech enhancement algorithms electronically – (PIs: Lansing, Zhao, and TBA).
2. Test the effectiveness of IHA devices in normal subjects and hearing-impaired subjects of various etiologies and different age groups – (PIs: Bilger, Gupta, and TBA).

Upon completion of the design and testing of the real time system, we will focus on microminiturizing the IHA devices for ultimate feasibility as practical hearing assistive devices. The miniaturization portion of the project would include designing radio-transmission remote-control-system, packaging, battery, etc. Additional expertise will be recruited from the Microelectronics Center (e.g., Steve Kang) to complete this phase of the project.

List of investigators (collaborators):

Robert Bilger	Professor of Speech and Hearing
Albert S. Feng	Professor of Beckman Institute & Molecular & Integrative Physiology
Prahlad Gupta	Beckman Fellow (starting date in the BI: January 1997) - He will be a part-time researcher on this project
Charissa Lansing	Assistant Professor of Beckman Institute & Speech and Hearing
Chen Liu	Postdoctoral Fellow at the Beckman Institute
William O'Brien	Professor of Electrical and Computer Engineering
Bruce Wheeler	Associate Professor of Beckman Institute & Electrical and Computer Eng.
Yun-Xin Zhao	Assistant Professor of Beckman Institute & Electrical and Computer Eng.
TBAs	Graduate Research Assistants - To be announced

Budget:

Postdoctoral salary for Dr. Chen Liu for two years (8/96 to 8/98)	$ 52,000
Salary for one TBA graduate research assistant for two years (8/96 to 8/98)	$ 28,000
Salary for one TBA graduate research assistant for one year (8/97 to 8/98)	$ 14,000
Electronic parts and supplies	$ 6,000
Total amount requested for two years	$100,000

Budget Justifications:

This collaborative project is being pursued at the Beckman Institute under the support of Jiri Jonas. His support covers the initial purchase of testing equipment and the first year salary for Dr. Chen Liu, a postdoc having extensive experience with the design of hearing aid devices. Sound generation and testing equipment, and a workstation along with appropriate interface hardware and software, were purchased earlier this year. These are installed in a sound isolated booth located in assigned laboratory space at the Beckman Institute.

Design of hearing prosthetic devices is of high priority for the National Institute for Deafness and Other Communicative Disorders. When we have pilot data to demonstrate the feasibility of the project the chance of receiving a research grant is high.

G

Subject: SBIR and STTR
Date: Monday, February 12, 1996 10:40AM

Hi, Lynn. Congratulations for the new addition to your family! I am pleased to hear from Amy that both the baby and the mother are doing fine.

I tried to call you this morning but you were away from your desk. So I decided to communicate by email. I write to ask if you would be so kind to send me the brochures and application packages for SBIR and STTR.

I informed Amy during the ARO meeting in Tampa that I am leading a group of researchers at the Beckman Institute to develop hearing aid devices that can help the hearing-impaired subjects to hear in complex acoustic environments with multi talkers (such as a cocktail party). The algorithm is neurally inspired and is looking very promising at least in our computer simulation. Our efforts are presently supported by the Beckman Institute (We did not think we had a fair chance if we were to compete with the high power groups of researchers before we got our feet wet and established our ground). The projest has progressed exceedingly well; the simulation results show that a speech signal can be faithfully extracted even when a babble noise of equal intensity (S/N of 0 dB) is broadcasted from an off-axis. We are carrying out a detailed evaluation of the algorithm. With our initial rapid success, we are encouraged that this direction is a great way to go for developing devices that can function well in complex noisy conditions.

To further this research, our team would benefit greatly if we can obtain financial support from the NIDCD. Amy indicated that such research projects have high priorities, and the success can bolster the credibility and publicity of the Institute, and that NIDCD would be interested in supporting promising research in this area. She suggested that we explore the SBIR (or STTR which, according to Bruce Wheeler, one of my collaborators, may be a better way to go) to reduce the risk of premature extensive exposure of the project before it has a chance to prove its overall capability. I believe we have a strong team: Bruce Wheeler, William O'Brien and Yun-Xin Zhao (three faculty members from the Electrical and Computer Engineering department specializing in signal processing and speech recognition), Charrissa Lansing and Bob Bilger (two professors from Speech and Hearing), myself and a postdoc who specialized in hearing aid research during his PhD dissertation. Additionally, once the algorithm has been proven and tested, several faculty from our renowned Microelectronics Center have expressed interest and support for miniturizing the devices.

I would appreciate it if you can forward to me the brochures and application packages for SBIR and STTR.

With best regards,

Al